

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 14:23:32 ; Search time 82 Seconds
(without alignments)
135.354 Million cell updates/sec

Title: US-09-825-489-3
Perfect score: 20
Sequence: 1 ggtccatcactcatgtgatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 800 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	100.0	822	4 US-09-651-656-8	Sequence 8, Appl1
C 2	20	100.0	822	4 US-09-650-855-8	Sequence 8, Appl1
C 3	16.8	84.0	14273	4 US-08-961-527-40	Sequence 40, Appl1
C 4	15.8	79.0	614	3 US-09-385-982-62	Sequence 62, Appl1
C 5	15.8	79.0	686	4 US-09-023-655-1225	Sequence 1225, App
C 6	15.8	79.0	801	4 US-09-543-681A-137	Sequence 137, App
C 7	15.8	79.0	1176	4 US-09-543-681A-1773	Sequence 1773, App
C 8	15.4	77.0	675	4 US-09-134-001C-388	Sequence 388, App
C 9	15.2	76.0	330	4 US-08-956-171E-1137	Sequence 1137, App
C 10	15.2	76.0	1419	4 US-09-489-039A-5224	Sequence 5224, App
C 11	15.2	76.0	4884	4 US-09-328-352-2478	Sequence 2478, App
C 12	15.2	76.0	14311	3 US-08-646-695-1	Sequence 1, Appl1
C 13	15.2	76.0	14311	3 US-08-646-695-7	Sequence 7, Appl1
C 14	15.2	76.0	14311	5 PCT-US96-06053-1	Sequence 1, Appl1
C 15	15.2	76.0	14311	5 PCT-US96-06053-7	Sequence 7, Appl1
C 16	15.2	76.0	36412	4 US-08-311-731A-132	Sequence 132, App
C 17	15.2	76.0	1230025	4 US-09-198-452A-1	Sequence 1, Appl1
C 18	15.2	76.0	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
C 19	15.2	76.0	4411529	3 US-09-103-840A-1	Sequence 1, Appl1
C 20	14.8	74.0	315	4 US-09-328-352-1128	Sequence 1128, App
C 21	14.8	74.0	348	4 US-09-328-352-1227	Sequence 1227, App
C 22	14.8	74.0	387	4 US-09-621-976-13741	Sequence 3741, App
C 23	14.8	74.0	546	4 US-09-328-352-992	Sequence 992, App
C 24	14.4	72.0	2244	4 US-09-489-039A-911	Sequence 911, App
C 25	14.4	72.0	2310	4 US-09-489-039A-953	Sequence 953, App
C 26	14.4	72.0	2433	1 US-08-136-743B-40	Sequence 40, Appl1
C 27	14.4	72.0	2663	1 US-08-136-743B-3	Sequence 3, Appl1

28	14.4	72.0	50341	1 US-08-247-901C-1	Sequence 1, Appl1
29	14.4	72.0	50341	2 US-09-075-904-1	Sequence 1, Appl1
30	14.4	72.0	52297	4 US-09-426-436-1	Sequence 1, Appl1
31	14.4	72.0	52297	4 US-08-705-557-1	Sequence 1, Appl1
32	14.2	71.0	36	3 US-09-135-782-10	Sequence 10, Appl1
33	14.2	71.0	329	4 US-08-956-171E-986	Sequence 986, App
34	14.2	71.0	378	4 US-09-134-001C-2448	Sequence 2448, App
35	14.2	71.0	423	4 US-09-252-991A-11719	Sequence 11719, App
36	14.2	71.0	480	4 US-09-252-991A-13151	Sequence 13151, App
37	14.2	71.0	510	4 US-09-134-000C-244	Sequence 244, App
38	14.2	71.0	555	4 US-09-107-532A-735	Sequence 735, App
39	14.2	71.0	743	4 US-09-342-681C-11	Sequence 11, Appl1
40	14.2	71.0	909	4 US-09-252-991A-9153	Sequence 9153, App
41	14.2	71.0	1065	4 US-09-634-238-121	Sequence 121, App
42	14.2	71.0	1176	4 US-09-342-681C-14	Sequence 14, Appl1
43	14.2	71.0	1185	4 US-09-252-991A-11577	Sequence 11577, App
44	14.2	71.0	1225	1 US-08-286-020-1	Sequence 1, Appl1
45	14.2	71.0	1225	1 US-08-603-919-1	Sequence 1, Appl1
46	14.2	71.0	1315	2 US-08-343-101A-8	Sequence 8, Appl1
47	14.2	71.0	1315	3 US-09-183-688-8	Sequence 8, Appl1
48	14.2	71.0	1315	4 US-09-519-991A-11863	Sequence 8, Appl1
49	14.2	71.0	1386	4 US-09-252-991A-11863	Sequence 8, Appl1
50	14.2	71.0	1398	4 US-09-489-039A-3739	Sequence 3739, App
51	14.2	71.0	1419	4 US-09-252-991A-11636	Sequence 11636, App
52	14.2	71.0	1574	4 US-09-342-681C-1	Sequence 1, Appl1
53	14.2	71.0	1651	4 US-09-533-029-59	Sequence 59, Appl1
54	14.2	71.0	1654	4 US-09-634-238-16	Sequence 16, Appl1
55	14.2	71.0	1707	4 US-09-134-001C-931	Sequence 931, App
56	14.2	71.0	1845	4 US-09-307-973A-1	Sequence 1, Appl1
57	14.2	71.0	1875	4 US-09-252-991A-8812	Sequence 8812, App
58	14.2	71.0	1901	5 PCT-US93-05000-32	Sequence 32, Appl1
59	14.2	71.0	1954	4 US-09-023-655-286	Sequence 286, App
60	14.2	71.0	1989	4 US-09-377-466B-5	Sequence 5, Appl1
61	14.2	71.0	1989	4 US-09-377-466B-7	Sequence 7, Appl1
62	14.2	71.0	2000	3 US-08-996-441B-99	Sequence 99, Appl1
63	14.2	71.0	2000	3 US-08-993-722A-99	Sequence 99, Appl1
64	14.2	71.0	2000	3 US-08-993-722A-99	Sequence 99, Appl1
65	14.2	71.0	2000	3 US-08-993-722A-99	Sequence 99, Appl1
66	14.2	71.0	2000	4 US-09-427-770-99	Sequence 99, Appl1
67	14.2	71.0	2000	4 US-09-427-770-99	Sequence 99, Appl1
68	14.2	71.0	2022	2 US-08-464-517-32	Sequence 32, Appl1
69	14.2	71.0	2022	2 US-08-246-361A-32	Sequence 32, Appl1
70	14.2	71.0	2022	2 US-08-463-772-32	Sequence 32, Appl1
71	14.2	71.0	2050	3 US-08-996-441B-101	Sequence 101, App
72	14.2	71.0	2050	3 US-08-993-722A-101	Sequence 101, App
73	14.2	71.0	2050	3 US-08-993-722A-101	Sequence 101, App
74	14.2	71.0	2050	3 US-08-993-722A-101	Sequence 101, App
75	14.2	71.0	2050	4 US-09-427-770-101	Sequence 101, App
76	14.2	71.0	2050	4 US-09-427-770-101	Sequence 101, App
77	14.2	71.0	2152	1 US-08-188-582-17	Sequence 17, Appl1
78	14.2	71.0	2152	1 US-08-646-715-17	Sequence 17, Appl1
79	14.2	71.0	2211	3 US-09-318-448-26	Sequence 26, Appl1
80	14.2	71.0	2211	4 US-09-347-878-2	Sequence 2, Appl1
81	14.2	71.0	2211	4 US-09-546-013-4	Sequence 4, Appl1
82	14.2	71.0	2259	1 US-08-420-235B-20	Sequence 20, Appl1
83	14.2	71.0	2259	3 US-08-793-624-20	Sequence 20, Appl1
84	14.2	71.0	2259	5 PCT-US95-10194-20	Sequence 20, Appl1
85	14.2	71.0	2304	4 US-09-252-991A-11802	Sequence 11802, App
86	14.2	71.0	2344	4 US-07-695-472B-35	Sequence 35, Appl1
87	14.2	71.0	2344	4 US-09-106-375-35	Sequence 35, Appl1
88	14.2	71.0	2347	4 US-08-250-740-32	Sequence 32, Appl1
89	14.2	71.0	2347	1 US-07-695-472B-34	Sequence 1, Appl1
90	14.2	71.0	2347	1 US-07-695-472B-34	Sequence 34, Appl1
91	14.2	71.0	2347	1 US-07-695-472B-36	Sequence 36, Appl1
92	14.2	71.0	2347	4 US-09-106-375-34	Sequence 34, Appl1
93	14.2	71.0	2347	4 US-09-106-375-36	Sequence 36, Appl1
94	14.2	71.0	2373	4 US-09-023-655-1011	Sequence 1011, App
95	14.2	71.0	2451	1 US-07-731-157A-3	Sequence 3, Appl1
96	14.2	71.0	2451	1 US-08-229-444B-1	Sequence 1, Appl1
97	14.2	71.0	2451	2 US-08-541-780-3	Sequence 3, Appl1
98	14.2	71.0	2517	4 US-09-252-991A-9238	Sequence 9238, App
99	14.2	71.0	2549	1 US-08-470-720-2	Sequence 2, Appl1
100	14.2	71.0	2549	1 US-08-470-720-2	Sequence 2, Appl1

101	14.2	71.0	2706	4	US-09-134-001C-2552	Sequence 2552, Ap	C 174	13.8	69.0	3797	1	US-08-789-449-1	Sequence 1, Appl
C 102	14.2	71.0	3043	4	US-09-049-698-16	Sequence 16, Appl	C 175	13.8	69.0	11707	3	US-09-136-574A-1	Sequence 1, Appl
C 103	14.2	71.0	3044	4	US-09-377-466B-38	Sequence 38, Appl	C 176	13.8	69.0	2866	4	US-09-596-002-15	Sequence 15, Appl
C 104	14.2	71.0	3181	4	US-09-049-698-18	Sequence 18, Appl	C 177	13.8	69.0	36682	3	US-08-943-731-2	Sequence 2, Appl
C 105	14.2	71.0	3198	4	US-09-345-236B-83	Sequence 83, Appl	C 178	13.8	69.0	43676	3	US-09-356-952-12	Sequence 12, Appl
C 106	14.2	71.0	3198	4	US-09-345-236B-87	Sequence 87, Appl	C 179	13.8	69.0	43512	4	US-09-671-317-485	Sequence 485, App
C 107	14.2	71.0	3198	4	US-09-345-236B-90	Sequence 90, Appl	C 180	13.8	69.0	63588	4	US-09-873-404-3	Sequence 3, Appl
C 108	14.2	71.0	3198	4	US-09-345-236B-92	Sequence 92, Appl	C 181	13.8	69.0	1230025	4	US-09-198-452A-1	Sequence 1, Appl
C 109	14.2	71.0	3362	2	US-08-678-039A-41	Sequence 41, Appl	C 182	13.8	69.0	1830121	4	US-09-557-884-1	Sequence 1, Appl
C 110	14.2	71.0	3455	4	US-09-377-466B-36	Sequence 36, Appl	C 183	13.8	69.0	1830121	4	US-09-643-990A-1	Sequence 1, Appl
C 111	14.2	71.0	3518	4	US-09-412-210-2	Sequence 2, Appl	C 184	13.6	68.0	115	4	US-09-494-921-20	Sequence 20, Appl
C 112	14.2	71.0	3518	4	US-10-121-911A-2	Sequence 2, Appl	C 185	13.6	68.0	272	4	US-09-016-434-95	Sequence 495, App
C 113	14.2	71.0	3500	4	US-09-134-000C-2137	Sequence 2137, Ap	C 186	13.6	68.0	289	4	US-09-313-294A-2910	Sequence 2910, Ap
C 114	14.2	71.0	3754	4	US-09-377-466B-15	Sequence 15, Appl	C 187	13.6	68.0	315	4	US-09-107-532A-2050	Sequence 2050, Ap
C 115	14.2	71.0	4008	4	US-09-134-001C-879	Sequence 879, App	C 188	13.6	68.0	336	4	US-09-328-352-3353	Sequence 3353, Ap
C 116	14.2	71.0	4149	4	US-09-377-466B-13	Sequence 13, Appl	C 189	13.6	68.0	400	4	US-08-956-111E-3987	Sequence 3987, Ap
C 117	14.2	71.0	4223	4	US-09-620-312D-857	Sequence 857, App	C 190	13.6	68.0	406	4	US-08-956-111E-3944	Sequence 3944, Ap
C 118	14.2	71.0	4381	4	US-09-347-878-19	Sequence 19, Appl	C 191	13.6	68.0	409	4	US-09-621-976-3157	Sequence 3157, Ap
C 119	14.2	71.0	4741	1	US-07-695-472B-4	Sequence 4, Appl	C 192	13.6	68.0	410	4	US-09-494-921-6	Sequence 6, Appl
C 120	14.2	71.0	4741	1	US-09-106-375-4	Sequence 4, Appl	C 193	13.6	68.0	430	4	US-09-621-976-3161	Sequence 3161, Ap
C 121	14.2	71.0	4742	1	US-08-250-740-35	Sequence 35, Appl	C 194	13.6	68.0	451	4	US-09-621-976-3161	Sequence 3161, Ap
C 122	14.2	71.0	4748	3	US-09-331-581-1	Sequence 1, Appl	C 195	13.6	68.0	456	4	US-09-107-532A-2457	Sequence 2457, Ap
C 123	14.2	71.0	4851	4	US-09-221-017B-999	Sequence 999, App	C 196	13.6	68.0	473	4	US-09-621-976-856	Sequence 856, App
C 124	14.2	71.0	4964	1	US-08-470-720-5	Sequence 5, Appl	C 197	13.6	68.0	494	4	US-09-621-976-3160	Sequence 3160, Ap
C 125	14.2	71.0	6157	4	US-08-956-171E-184	Sequence 184, Appl	C 198	13.6	68.0	528	4	US-09-621-976-3159	Sequence 3159, Ap
C 126	14.2	71.0	6464	4	US-09-221-017B-168	Sequence 168, App	C 199	13.6	68.0	593	3	US-08-961-083-107	Sequence 107, App
C 127	14.2	71.0	7286	3	US-09-331-581-3	Sequence 3, Appl	C 200	13.6	68.0	593	4	US-09-536-784-107	Sequence 107, App
C 128	14.2	71.0	7938	3	US-09-331-581-14	Sequence 14, Appl	C 201	13.6	68.0	642	4	US-09-831-381-1086	Sequence 1086, App
C 129	14.2	71.0	11340	4	US-08-961-527-147	Sequence 147, App	C 202	13.6	68.0	680	4	US-09-227-357-144	Sequence 144, App
C 130	14.2	71.0	20710	1	US-08-420-235B-1	Sequence 1, Appl	C 203	13.6	68.0	751	4	US-09-621-976-3052	Sequence 3052, Ap
C 131	14.2	71.0	20710	3	US-08-793-624-1	Sequence 1, Appl	C 204	13.6	68.0	876	2	US-08-928-284-1	Sequence 1, Appl
C 132	14.2	71.0	20710	5	PCR-US85-10194-1	Sequence 1, Appl	C 205	13.6	68.0	975	3	US-09-365-150-3	Sequence 3, Appl
C 133	14.2	71.0	35100	3	US-08-770-379-18	Sequence 18, Appl	C 206	13.6	68.0	975	3	US-09-365-150-4	Sequence 4, Appl
C 134	14.2	71.0	35100	3	US-08-757-669A-18	Sequence 18, Appl	C 207	13.6	68.0	1158	4	US-09-491-577-41	Sequence 41, Appl
C 135	14.2	71.0	35100	4	US-09-230-371A-18	Sequence 18, Appl	C 208	13.6	68.0	1197	3	US-09-383-586-5	Sequence 5, Appl
C 136	14.2	71.0	36063	4	US-08-311-731A-140	Sequence 140, App	C 209	13.6	68.0	1197	4	US-09-494-921-1	Sequence 1, Appl
C 137	14.2	71.0	36241	4	US-08-311-731A-134	Sequence 134, App	C 210	13.6	68.0	1197	4	US-09-494-921-18	Sequence 28, Appl
C 138	14.2	71.0	42157	4	US-08-311-731A-126	Sequence 126, App	C 211	13.6	68.0	1197	4	US-09-494-921-29	Sequence 29, Appl
C 139	14.2	71.0	580073	4	US-08-545-528D-1	Sequence 1, Appl	C 212	13.6	68.0	1200	4	US-09-494-921-30	Sequence 30, Appl
C 140	14.2	71.0	580073	4	US-08-545-528D-1	Sequence 1, Appl	C 213	13.6	68.0	1202	4	US-09-494-921-3	Sequence 3, Appl
C 141	14.2	71.0	1664976	4	US-08-916-421B-1	Sequence 1, Appl	C 214	13.6	68.0	1209	4	US-09-489-039A-344	Sequence 344, App
C 142	14.2	71.0	435	4	US-09-328-352-1934	Sequence 1934, Ap	C 215	13.6	68.0	1216	3	US-08-838-151A-62	Sequence 62, Appl
C 143	14.2	71.0	2079	3	US-09-489-847-25	Sequence 25, Appl	C 216	13.6	68.0	1247	4	US-09-740-035-1	Sequence 1, Appl
C 144	13.8	69.0	279	3	US-09-328-111-613	Sequence 613, App	C 217	13.6	68.0	1260	3	US-09-008-979A-3	Sequence 3, Appl
C 145	13.8	69.0	426	1	US-08-470-179-69	Sequence 69, Appl	C 218	13.6	68.0	1260	3	US-09-460-618-3	Sequence 3, Appl
C 146	13.8	69.0	466	4	US-09-621-976-11011	Sequence 11011, A	C 219	13.6	68.0	1260	4	US-09-310-235B-3	Sequence 3, Appl
C 147	13.8	69.0	486	4	US-09-543-681A-2354	Sequence 2354, Ap	C 220	13.6	68.0	1263	4	US-09-252-981A-10082	Sequence 10082, A
C 148	13.8	69.0	633	3	US-09-352-990-15	Sequence 15, Appl	C 221	13.6	68.0	1302	4	US-09-170-46D-47	Sequence 47, Appl
C 149	13.8	69.0	638	4	US-09-669-751-66	Sequence 66, Appl	C 222	13.6	68.0	1302	4	US-09-170-46D-189	Sequence 189, App
C 150	13.8	69.0	741	4	US-09-489-039A-3825	Sequence 3825, Ap	C 223	13.6	68.0	1320	2	US-08-641-038A-1	Sequence 1, Appl
C 151	13.8	69.0	766	2	US-08-763-121-2	Sequence 2, Appl	C 224	13.6	68.0	1320	2	US-09-059-178-1	Sequence 1, Appl
C 152	13.8	69.0	766	4	US-09-216-066-2	Sequence 2, Appl	C 225	13.6	68.0	1329	4	US-09-328-352-2751	Sequence 2751, Ap
C 153	13.8	69.0	1107	1	US-08-844-055-1	Sequence 1, Appl	C 226	13.6	68.0	1332	4	US-09-134-001C-311	Sequence 311, App
C 154	13.8	69.0	1107	3	US-09-006-849-1	Sequence 1, Appl	C 227	13.6	68.0	1335	4	US-09-540-226-1425	Sequence 1425, Ap
C 155	13.8	69.0	1119	4	US-08-773-423-6	Sequence 8546, Ap	C 228	13.6	68.0	1359	4	US-09-489-039A-5387	Sequence 5387, Ap
C 156	13.8	69.0	1175	2	US-08-906-744A-1	Sequence 6, Appl	C 229	13.6	68.0	1359	2	US-08-642-541-1	Sequence 1, Appl
C 157	13.8	69.0	1290	3	US-09-093-134-1	Sequence 1, Appl	C 230	13.6	68.0	1369	2	US-08-642-541-3	Sequence 3, Appl
C 158	13.8	69.0	1290	3	US-09-093-134-1	Sequence 1, Appl	C 231	13.6	68.0	1369	3	US-09-260-889-1	Sequence 1, Appl
C 159	13.8	69.0	1458	4	US-09-482-273-14	Sequence 14, Appl	C 232	13.6	68.0	1369	3	US-09-260-889-3	Sequence 3, Appl
C 160	13.8	69.0	1464	4	US-09-328-352-3361	Sequence 3361, Ap	C 233	13.6	68.0	1369	4	US-09-347-878-11	Sequence 11, Appl
C 161	13.8	69.0	1550	3	US-08-943-731-70	Sequence 70, Appl	C 234	13.6	68.0	1369	4	US-09-479-275-6	Sequence 6, Appl
C 162	13.8	69.0	1569	4	US-09-328-352-2714	Sequence 2714, Ap	C 235	13.6	68.0	1380	4	US-09-134-001C-2629	Sequence 2629, Ap
C 163	13.8	69.0	1740	4	US-09-489-039A-5163	Sequence 5163, Ap	C 236	13.6	68.0	1457	4	US-09-419-679-9	Sequence 9, Appl
C 164	13.8	69.0	2029	3	US-09-136-574A-46	Sequence 46, Appl	C 237	13.6	68.0	1479	4	US-09-252-991A-6291	Sequence 6291, Ap
C 165	13.8	69.0	2184	4	US-09-543-681A-2518	Sequence 2518, Ap	C 238	13.6	68.0	1512	4	US-09-107-532A-54	Sequence 54, Appl
C 166	13.8	69.0	3311	2	US-08-239-276-10	Sequence 10, Appl	C 239	13.6	68.0	1662	4	US-09-252-991A-6018	Sequence 6018, Ap
C 167	13.8	69.0	3311	3	US-08-468-579B-10	Sequence 10, Appl	C 240	13.6	68.0	1721	4	US-09-197-679A-1	Sequence 1, Appl
C 168	13.8	69.0	3311	3	US-08-468-577B-10	Sequence 10, Appl	C 241	13.6	68.0	1731	4	US-09-506-066E-3	Sequence 3, Appl
C 169	13.8	69.0	3606	4	US-08-956-171E-43	Sequence 43, Appl	C 242	13.6	68.0	1796	1	US-07-816-283-11	Sequence 11, Appl
C 170	13.8	69.0	3613	4	US-08-514-213A-1	Sequence 5, Appl	C 243	13.6	68.0	1796	1	US-08-417-103-11	Sequence 11, Appl
C 171	13.8	69.0	3637	4	US-09-566-921-5	Sequence 5, Appl	C 244	13.6	68.0	1855	4	US-09-403-752A-19	Sequence 19, Appl
C 172	13.8	69.0	3797	1	US-07-915-203-1	Sequence 1, Appl	C 245	13.6	68.0	2000	1	US-07-923-739-1	Sequence 1, Appl
C 173	13.8	69.0	3797	1	US-08-872-887-1	Sequence 1, Appl	C 246	13.6	68.0	2002	2	US-08-715-202A-3	Sequence 3, Appl

C 247	13.6	68.0	2002	4	US-09-328-775-3	Sequence 3, Appl1	320	13.2	66.0	185	4	US-09-313-294A-4007	Sequence 4007, Ap
C 248	13.6	68.0	2002	4	US-09-994-177-3	Sequence 3, Appl1	C 321	13.2	66.0	238	1	US-08-435-684A-56	Sequence 56, Appl1
C 249	13.6	68.0	2045	4	US-08-753-750B-5	Sequence 5	C 322	13.2	66.0	238	2	US-08-934-877A-56	Sequence 56, Appl1
C 250	13.6	68.0	2211	4	US-09-016-434-1150	Sequence 1150, Ap	C 323	13.2	66.0	238	3	US-08-871-678C-56	Sequence 56, Appl1
C 251	13.6	68.0	2238	2	US-08-919-624-2	Sequence 2, Appl1	C 324	13.2	66.0	289	4	US-09-313-294A-1390	Sequence 1390, Ap
C 252	13.6	68.0	2259	1	US-08-145-681-5	Sequence 5, Appl1	C 325	13.2	66.0	306	4	US-09-543-681A-121	Sequence 121, Ap
C 253	13.6	68.0	2259	1	US-08-453-703-5	Sequence 5, Appl1	C 326	13.2	66.0	327	3	US-08-766-355-9	Sequence 9, Appl1
C 254	13.6	68.0	2259	2	US-08-456-106-5	Sequence 5, Appl1	C 327	13.2	66.0	327	3	US-09-003-198A-9	Sequence 9, Appl1
C 255	13.6	68.0	2259	3	US-08-456-108-5	Sequence 5, Appl1	C 328	13.2	66.0	327	3	US-09-428-805-9	Sequence 9, Appl1
C 256	13.6	68.0	2259	3	US-09-265-577-5	Sequence 5, Appl1	C 329	13.2	66.0	352	3	US-08-961-083-127	Sequence 127, Ap
C 257	13.6	68.0	2259	4	US-09-633-739-5	Sequence 5, Appl1	C 330	13.2	66.0	352	4	US-09-536-784-127	Sequence 127, Ap
C 258	13.6	68.0	2278	4	US-09-276-438-10	Sequence 10, Appl1	C 331	13.2	66.0	384	4	US-09-543-681A-1889	Sequence 1889, Ap
C 259	13.6	68.0	2587	2	US-08-326-286-3	Sequence 3, Appl1	C 332	13.2	66.0	396	4	US-09-621-976-1322	Sequence 8322, Ap
C 260	13.6	68.0	2817	4	US-08-956-171E-295	Sequence 295, Ap	C 333	13.2	66.0	430	4	US-09-621-976-1254	Sequence 8254, Ap
C 261	13.6	68.0	2896	2	US-08-709-923-1	Sequence 1, Appl1	C 334	13.2	66.0	445	4	US-09-621-976-15305	Sequence 15305, A
C 262	13.6	68.0	3375	1	US-08-381-931B-1	Sequence 1, Appl1	C 335	13.2	66.0	482	4	US-09-621-976-115	Sequence 315, Ap
C 263	13.6	68.0	3466	4	US-09-782-906-1	Sequence 1, Appl1	C 336	13.2	66.0	490	4	US-09-003-198A-19	Sequence 19, Appl1
C 264	13.6	68.0	3784	1	US-07-623-033-1	Sequence 1, Appl1	C 337	13.2	66.0	495	4	US-09-543-681A-2833	Sequence 2833, Ap
C 265	13.6	68.0	3993	4	US-09-543-681A-3885	Sequence 3885, Ap	C 338	13.2	66.0	507	3	US-08-766-355-10	Sequence 10, Appl1
C 266	13.6	68.0	4081	4	US-08-930-055A-1	Sequence 1, Appl1	C 339	13.2	66.0	507	4	US-09-003-198A-10	Sequence 10, Appl1
C 267	13.6	68.0	4617	4	US-08-543-681A-1157	Sequence 1157, Ap	C 340	13.2	66.0	507	4	US-09-428-805-10	Sequence 10, Appl1
C 268	13.6	68.0	4761	4	US-08-920-812-7	Sequence 7, Appl1	C 341	13.2	66.0	510	4	US-09-134-001C-1698	Sequence 1698, Ap
C 269	13.6	68.0	5024	1	US-08-920-812-7	Sequence 7, Appl1	C 342	13.2	66.0	522	4	US-09-598-401C-61	Sequence 61, Appl1
C 270	13.6	68.0	5024	1	US-08-920-812-7	Sequence 7, Appl1	C 343	13.2	66.0	537	4	US-09-621-976-10463	Sequence 10563, A
C 271	13.6	68.0	5024	1	US-08-921-177-7	Sequence 7, Appl1	C 344	13.2	66.0	538	4	US-09-621-976-1523	Sequence 2523, Ap
C 272	13.6	68.0	5024	1	US-08-362-577C-7	Sequence 7, Appl1	C 345	13.2	66.0	600	4	US-09-621-976-1524	Sequence 2524, Ap
C 273	13.6	68.0	5024	2	US-08-920-828-7	Sequence 7, Appl1	C 346	13.2	66.0	640	3	US-09-328-111-792	Sequence 792, Ap
C 274	13.6	68.0	5109	4	US-08-930-055A-2	Sequence 2, Appl1	C 347	13.2	66.0	651	4	US-09-134-001C-781	Sequence 781, Ap
C 275	13.6	68.0	5455	1	US-08-342-930-1	Sequence 1, Appl1	C 348	13.2	66.0	688	4	US-09-972-800A-70	Sequence 70, Appl1
C 276	13.6	68.0	6102	4	US-09-976-594-682	Sequence 682, Ap	C 349	13.2	66.0	767	4	US-09-921-017B-28	Sequence 28, Appl1
C 277	13.6	68.0	6102	4	US-08-961-527-167	Sequence 167, Ap	C 350	13.2	66.0	813	4	US-09-872-733A-19	Sequence 19, Appl1
C 278	13.6	68.0	29555	4	US-08-956-171E-206	Sequence 206, Ap	C 351	13.2	66.0	846	1	US-07-672-304-2	Sequence 2, Appl1
C 279	13.6	68.0	29555	4	US-08-956-171E-206	Sequence 206, Ap	C 352	13.2	66.0	870	4	US-09-328-352-1901	Sequence 1901, Ap
C 280	13.6	68.0	36800	4	US-08-311-731A-139	Sequence 139, Ap	C 353	13.2	66.0	876	4	US-09-134-000C-860	Sequence 860, Ap
C 281	13.6	68.0	37769	4	US-08-311-731A-23	Sequence 23, Appl1	C 354	13.2	66.0	885	4	US-09-543-681A-1746	Sequence 1746, Ap
C 282	13.6	68.0	56516	2	US-08-996-306-1	Sequence 1, Appl1	C 355	13.2	66.0	887	4	US-09-247-155-150	Sequence 150, Ap
C 283	13.6	68.0	56516	3	US-09-338-907-1	Sequence 1, Appl1	C 356	13.2	66.0	899	4	US-09-221-017B-1060	Sequence 1060, Ap
C 284	13.6	68.0	56516	4	US-09-218-207-1	Sequence 1, Appl1	C 357	13.2	66.0	914	4	US-09-673-395A-114	Sequence 114, Ap
C 285	13.6	68.0	56520	3	US-09-338-907-179	Sequence 179, Ap	C 358	13.2	66.0	921	4	US-09-489-039A-1033	Sequence 1033, Ap
C 286	13.6	68.0	56520	4	US-09-218-207-179	Sequence 179, Ap	C 359	13.2	66.0	923	4	US-09-221-017B-252	Sequence 252, Ap
C 287	13.6	68.0	66986	4	US-09-596-002-29	Sequence 29, Appl1	C 360	13.2	66.0	933	4	US-09-107-532A-558	Sequence 558, Ap
C 288	13.6	68.0	392000	4	US-10-027-983-11	Sequence 11, Appl1	C 361	13.2	66.0	951	4	US-09-543-681A-2787	Sequence 2787, Ap
C 289	13.6	68.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl1	C 362	13.2	66.0	969	4	US-09-543-681A-3696	Sequence 3696, Ap
C 290	13.6	68.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl1	C 363	13.2	66.0	1002	4	US-09-641-638-883	Sequence 583, Ap
C 291	13.4	67.0	266	4	US-09-313-294A-1855	Sequence 1855, Ap	C 364	13.2	66.0	1007	4	US-09-620-312D-817	Sequence 817, Ap
C 292	13.4	67.0	311	4	US-09-963-137-15	Sequence 15, Appl1	C 365	13.2	66.0	1021	4	US-08-936-171E-14	Sequence 14, Appl1
C 293	13.4	67.0	431	4	US-09-621-976-16186	Sequence 16186, A	C 366	13.2	66.0	1035	4	US-09-134-001C-687	Sequence 687, Ap
C 294	13.4	67.0	442	4	US-09-071-035-443	Sequence 443, Ap	C 367	13.2	66.0	1092	4	US-09-797-464A-10	Sequence 10, Appl1
C 295	13.4	67.0	453	4	US-08-817-441-64	Sequence 441, Ap	C 368	13.2	66.0	1113	4	US-09-540-236A-338	Sequence 338, Ap
C 296	13.4	67.0	606	4	US-09-071-035-441	Sequence 441, Ap	C 369	13.2	66.0	1116	3	US-08-916-443A-16	Sequence 16, Appl1
C 297	13.4	67.0	638	2	US-08-222-719-16	Sequence 16, Appl1	C 370	13.2	66.0	1142	4	US-08-936-165A-92	Sequence 92, Appl1
C 298	13.4	67.0	638	2	US-08-470-925-16	Sequence 16, Appl1	C 371	13.2	66.0	1233	3	US-08-853-839-1	Sequence 1, Appl1
C 299	13.4	67.0	638	5	US-08-471-613-16	Sequence 16, Appl1	C 372	13.2	66.0	1239	4	US-09-252-991A-12224	Sequence 12224, A
C 300	13.4	67.0	638	5	PCT-US93-10443-16	Sequence 16, Appl1	C 373	13.2	66.0	1242	4	US-09-543-681A-1921	Sequence 1921, Ap
C 301	13.4	67.0	690	4	US-09-134-000C-1299	Sequence 1299, Ap	C 374	13.2	66.0	1248	2	US-08-897-340-5	Sequence 5, Appl1
C 302	13.4	67.0	1123	4	US-08-956-171E-427	Sequence 427, Ap	C 375	13.2	66.0	1248	2	US-08-897-340-5	Sequence 5, Appl1
C 303	13.4	67.0	1473	3	US-08-907-740-6	Sequence 6, Appl1	C 376	13.2	66.0	1291	1	US-07-952-755-2	Sequence 2, Appl1
C 304	13.4	67.0	2616	4	US-09-107-532A-2573	Sequence 2573, Ap	C 377	13.2	66.0	1291	1	US-08-443-679-8	Sequence 215, Ap
C 305	13.4	67.0	2715	4	US-09-543-681A-513	Sequence 513, Ap	C 378	13.2	66.0	1308	4	US-09-205-268-215	Sequence 215, Ap
C 306	13.4	67.0	3001	4	US-09-539-333D-213	Sequence 213, Ap	C 379	13.2	66.0	1360	4	US-09-082-649B-80	Sequence 80, Appl1
C 307	13.4	67.0	3379	4	US-09-275-608-1	Sequence 1, Appl1	C 380	13.2	66.0	1378	2	US-08-776-585-2	Sequence 2, Appl1
C 308	13.4	67.0	4695	4	US-09-620-312D-379	Sequence 379, Ap	C 381	13.2	66.0	1400	2	US-08-776-585-1	Sequence 1, Appl1
C 309	13.4	67.0	6543	4	US-09-963-137-131	Sequence 131, Ap	C 382	13.2	66.0	1404	4	US-09-489-039A-4718	Sequence 4718, Ap
C 310	13.4	67.0	8608	4	US-09-620-312D-333	Sequence 333, Ap	C 383	13.2	66.0	1414	4	US-09-404-268B-9	Sequence 9, Appl1
C 311	13.4	67.0	24979	2	US-08-147-777-3	Sequence 3, Appl1	C 384	13.2	66.0	1429	2	US-08-244-205-6	Sequence 6, Appl1
C 312	13.4	67.0	24979	5	PCT-US93-03985-3	Sequence 3, Appl1	C 385	13.2	66.0	1429	5	PCT-US92-10284-6	Sequence 8, Appl1
C 313	13.4	67.0	40138	4	US-09-090-793-112	Sequence 12, Appl1	C 386	13.2	66.0	1429	5	PCT-US92-10284-6	Sequence 8, Appl1
C 314	13.4	67.0	40138	4	US-09-231-899-12	Sequence 12, Appl1	C 387	13.2	66.0	1455	4	US-09-489-039A-4800	Sequence 4800, Ap
C 315	13.4	67.0	51952	4	US-08-947-823-1	Sequence 1, Appl1	C 388	13.2	66.0	1457	4	US-09-673-359A-554	Sequence 554, Ap
C 316	13.4	67.0	536165	3	US-09-214-808-1	Sequence 1, Appl1	C 389	13.2	66.0	1479	4	US-09-489-039A-1847	Sequence 1847, Ap
C 317	13.4	67.0	786431	4	US-09-751-389-3	Sequence 3, Appl1	C 391	13.2	66.0	1482	3	US-08-996-441B-69	Sequence 69, Appl1
C 318	13.4	67.0	786431	4	US-09-972-800A-58	Sequence 58, Appl1	C 392	13.2	66.0	1482	3	US-08-993-722A-69	Sequence 69, Appl1

C 393	13.2	66.0	1482	3	US-08-993-170A-69	Sequence 69, Appl	C 466	13.2	66.0	1959	3	US-08-993-722A-47	Sequence 47, Appl
C 394	13.2	66.0	1482	3	US-08-993-775B-69	Sequence 69, Appl	C 467	13.2	66.0	1959	3	US-08-993-722A-49	Sequence 49, Appl
C 395	13.2	66.0	1482	3	US-09-427-770-69	Sequence 69, Appl	C 468	13.2	66.0	1959	3	US-08-993-722A-53	Sequence 53, Appl
C 396	13.2	66.0	1482	4	US-09-427-769-69	Sequence 69, Appl	C 469	13.2	66.0	1959	3	US-08-993-722A-63	Sequence 63, Appl
C 397	13.2	66.0	1486	1	US-07-672-304-5	Sequence 5, Appl	C 470	13.2	66.0	1959	3	US-08-993-722A-65	Sequence 65, Appl
C 398	13.2	66.0	1483	4	US-09-227-357-26	Sequence 26, Appl	C 471	13.2	66.0	1959	3	US-08-993-722A-67	Sequence 67, Appl
C 399	13.2	66.0	1600	3	US-08-917-299-36	Sequence 36, Appl	C 472	13.2	66.0	1959	3	US-08-993-722A-97	Sequence 97, Appl
C 400	13.2	66.0	1600	3	US-09-422-662-36	Sequence 36, Appl	C 473	13.2	66.0	1959	3	US-08-993-170A-1	Sequence 1, Appl
C 401	13.2	66.0	1600	4	US-09-730-763-36	Sequence 36, Appl	C 474	13.2	66.0	1959	3	US-08-993-170A-3	Sequence 3, Appl
C 402	13.2	66.0	1600	4	US-09-429-370-36	Sequence 36, Appl	C 475	13.2	66.0	1959	3	US-08-993-170A-5	Sequence 5, Appl
C 403	13.2	66.0	1644	4	US-09-489-039A-2502	Sequence 5202, Ap	C 476	13.2	66.0	1959	3	US-08-993-170A-7	Sequence 7, Appl
C 404	13.2	66.0	1647	4	US-09-543-681A-2459	Sequence 2459, Ap	C 477	13.2	66.0	1959	3	US-08-993-170A-9	Sequence 9, Appl
C 405	13.2	66.0	1682	4	US-09-620-312D-181	Sequence 181, App	C 478	13.2	66.0	1959	3	US-08-993-170A-11	Sequence 11, Appl
C 406	13.2	66.0	1691	3	US-08-917-299-3	Sequence 3, Appl	C 479	13.2	66.0	1959	3	US-08-993-170A-13	Sequence 13, Appl
C 407	13.2	66.0	1691	3	US-09-422-662-3	Sequence 3, Appl	C 480	13.2	66.0	1959	3	US-08-993-170A-17	Sequence 17, Appl
C 408	13.2	66.0	1691	4	US-09-730-763-3	Sequence 3, Appl	C 481	13.2	66.0	1959	3	US-08-993-170A-29	Sequence 29, Appl
C 409	13.2	66.0	1691	4	US-09-230-485-1	Sequence 1, Appl	C 482	13.2	66.0	1959	3	US-08-993-170A-39	Sequence 39, Appl
C 410	13.2	66.0	1738	4	US-09-574-873-9	Sequence 9, Appl	C 483	13.2	66.0	1959	3	US-08-993-170A-41	Sequence 41, Appl
C 411	13.2	66.0	1875	4	US-09-543-681A-1964	Sequence 1964, Ap	C 484	13.2	66.0	1959	3	US-08-993-170A-43	Sequence 43, Appl
C 412	13.2	66.0	1917	4	US-09-107-532A-265	Sequence 265, Ap	C 485	13.2	66.0	1959	3	US-08-993-170A-45	Sequence 45, Appl
C 413	13.2	66.0	1953	1	US-08-315-468-5	Sequence 5, Appl	C 486	13.2	66.0	1959	3	US-08-993-170A-47	Sequence 47, Appl
C 414	13.2	66.0	1953	3	US-07-941-650A-3	Sequence 3, Appl	C 487	13.2	66.0	1959	3	US-08-993-170A-49	Sequence 49, Appl
C 415	13.2	66.0	1956	3	US-08-996-441B-51	Sequence 51, Appl	C 488	13.2	66.0	1959	3	US-08-993-170A-53	Sequence 53, Appl
C 416	13.2	66.0	1956	3	US-08-996-441B-55	Sequence 55, Appl	C 489	13.2	66.0	1959	3	US-08-993-170A-65	Sequence 65, Appl
C 417	13.2	66.0	1956	3	US-08-996-441B-57	Sequence 57, Appl	C 490	13.2	66.0	1959	3	US-08-993-170A-67	Sequence 67, Appl
C 418	13.2	66.0	1956	3	US-08-993-722A-51	Sequence 51, Appl	C 491	13.2	66.0	1959	3	US-08-993-170A-69	Sequence 69, Appl
C 419	13.2	66.0	1956	3	US-08-993-722A-55	Sequence 55, Appl	C 492	13.2	66.0	1959	3	US-08-993-170A-97	Sequence 97, Appl
C 420	13.2	66.0	1956	3	US-08-993-722A-57	Sequence 57, Appl	C 493	13.2	66.0	1959	3	US-08-993-775B-3	Sequence 3, Appl
C 421	13.2	66.0	1956	3	US-08-993-170A-51	Sequence 51, Appl	C 494	13.2	66.0	1959	3	US-08-993-775B-5	Sequence 5, Appl
C 422	13.2	66.0	1956	3	US-08-993-170A-55	Sequence 55, Appl	C 495	13.2	66.0	1959	3	US-08-993-775B-7	Sequence 7, Appl
C 423	13.2	66.0	1956	3	US-08-993-170A-57	Sequence 57, Appl	C 496	13.2	66.0	1959	3	US-08-993-775B-9	Sequence 9, Appl
C 424	13.2	66.0	1956	3	US-08-993-775B-51	Sequence 51, Appl	C 497	13.2	66.0	1959	3	US-08-993-775B-11	Sequence 11, Appl
C 425	13.2	66.0	1956	3	US-08-993-775B-55	Sequence 55, Appl	C 498	13.2	66.0	1959	3	US-08-993-775B-13	Sequence 13, Appl
C 426	13.2	66.0	1956	4	US-09-427-770-51	Sequence 51, Appl	C 499	13.2	66.0	1959	3	US-08-993-775B-27	Sequence 27, Appl
C 427	13.2	66.0	1956	4	US-09-427-770-55	Sequence 55, Appl	C 500	13.2	66.0	1959	3	US-08-993-775B-29	Sequence 29, Appl
C 428	13.2	66.0	1956	4	US-09-427-770-57	Sequence 57, Appl	C 501	13.2	66.0	1959	3	US-08-993-775B-39	Sequence 39, Appl
C 429	13.2	66.0	1956	4	US-09-427-770-57	Sequence 57, Appl	C 502	13.2	66.0	1959	3	US-08-993-775B-43	Sequence 43, Appl
C 430	13.2	66.0	1956	4	US-09-427-769-51	Sequence 51, Appl	C 503	13.2	66.0	1959	3	US-08-993-775B-45	Sequence 45, Appl
C 431	13.2	66.0	1956	4	US-09-427-769-55	Sequence 55, Appl	C 504	13.2	66.0	1959	3	US-08-993-775B-47	Sequence 47, Appl
C 432	13.2	66.0	1956	4	US-09-427-769-57	Sequence 57, Appl	C 505	13.2	66.0	1959	3	US-08-993-775B-49	Sequence 49, Appl
C 433	13.2	66.0	1959	3	US-08-996-441B-1	Sequence 1, Appl	C 506	13.2	66.0	1959	3	US-08-993-775B-53	Sequence 53, Appl
C 434	13.2	66.0	1959	3	US-08-996-441B-3	Sequence 3, Appl	C 507	13.2	66.0	1959	3	US-08-993-775B-55	Sequence 55, Appl
C 435	13.2	66.0	1959	3	US-08-996-441B-5	Sequence 5, Appl	C 508	13.2	66.0	1959	3	US-08-993-775B-67	Sequence 67, Appl
C 436	13.2	66.0	1959	3	US-08-996-441B-7	Sequence 7, Appl	C 509	13.2	66.0	1959	3	US-08-993-775B-69	Sequence 69, Appl
C 437	13.2	66.0	1959	3	US-08-996-441B-11	Sequence 11, Appl	C 510	13.2	66.0	1959	3	US-08-993-775B-65	Sequence 65, Appl
C 438	13.2	66.0	1959	3	US-08-996-441B-13	Sequence 13, Appl	C 511	13.2	66.0	1959	3	US-08-993-775B-67	Sequence 67, Appl
C 439	13.2	66.0	1959	3	US-08-996-441B-27	Sequence 27, Appl	C 512	13.2	66.0	1959	3	US-08-993-775B-97	Sequence 97, Appl
C 440	13.2	66.0	1959	3	US-08-996-441B-29	Sequence 29, Appl	C 513	13.2	66.0	1959	4	US-09-377-466B-1	Sequence 1, Appl
C 441	13.2	66.0	1959	3	US-08-996-441B-29	Sequence 29, Appl	C 514	13.2	66.0	1959	4	US-09-377-466B-3	Sequence 3, Appl
C 442	13.2	66.0	1959	3	US-08-996-441B-39	Sequence 39, Appl	C 515	13.2	66.0	1959	4	US-09-427-770-1	Sequence 1, Appl
C 443	13.2	66.0	1959	3	US-08-996-441B-41	Sequence 41, Appl	C 516	13.2	66.0	1959	4	US-09-427-770-3	Sequence 3, Appl
C 444	13.2	66.0	1959	3	US-08-996-441B-45	Sequence 45, Appl	C 517	13.2	66.0	1959	4	US-09-427-770-5	Sequence 5, Appl
C 445	13.2	66.0	1959	3	US-08-996-441B-45	Sequence 45, Appl	C 518	13.2	66.0	1959	4	US-09-427-770-7	Sequence 7, Appl
C 446	13.2	66.0	1959	3	US-08-996-441B-47	Sequence 47, Appl	C 519	13.2	66.0	1959	4	US-09-427-770-9	Sequence 9, Appl
C 447	13.2	66.0	1959	3	US-08-996-441B-49	Sequence 49, Appl	C 520	13.2	66.0	1959	4	US-09-427-770-11	Sequence 11, Appl
C 448	13.2	66.0	1959	3	US-08-996-441B-53	Sequence 53, Appl	C 521	13.2	66.0	1959	4	US-09-427-770-13	Sequence 13, Appl
C 449	13.2	66.0	1959	3	US-08-996-441B-65	Sequence 65, Appl	C 522	13.2	66.0	1959	4	US-09-427-770-27	Sequence 27, Appl
C 450	13.2	66.0	1959	3	US-08-996-441B-65	Sequence 65, Appl	C 523	13.2	66.0	1959	4	US-09-427-770-29	Sequence 29, Appl
C 451	13.2	66.0	1959	3	US-08-996-441B-67	Sequence 67, Appl	C 524	13.2	66.0	1959	4	US-09-427-770-39	Sequence 39, Appl
C 452	13.2	66.0	1959	3	US-08-996-441B-97	Sequence 97, Appl	C 525	13.2	66.0	1959	4	US-09-427-770-41	Sequence 41, Appl
C 453	13.2	66.0	1959	3	US-08-993-722A-1	Sequence 1, Appl	C 526	13.2	66.0	1959	4	US-09-427-770-43	Sequence 43, Appl
C 454	13.2	66.0	1959	3	US-08-993-722A-3	Sequence 3, Appl	C 527	13.2	66.0	1959	4	US-09-427-770-45	Sequence 45, Appl
C 455	13.2	66.0	1959	3	US-08-993-722A-5	Sequence 5, Appl	C 528	13.2	66.0	1959	4	US-09-427-770-47	Sequence 47, Appl
C 456	13.2	66.0	1959	3	US-08-993-722A-7	Sequence 7, Appl	C 529	13.2	66.0	1959	4	US-09-427-770-49	Sequence 49, Appl
C 457	13.2	66.0	1959	3	US-08-993-722A-9	Sequence 9, Appl	C 530	13.2	66.0	1959	4	US-09-427-770-53	Sequence 53, Appl
C 458	13.2	66.0	1959	3	US-08-993-722A-11	Sequence 11, Appl	C 531	13.2	66.0	1959	4	US-09-427-770-63	Sequence 63, Appl
C 459	13.2	66.0	1959	3	US-08-993-722A-17	Sequence 17, Appl	C 532	13.2	66.0	1959	4	US-09-427-770-65	Sequence 65, Appl
C 460	13.2	66.0	1959	3	US-08-993-722A-23	Sequence 23, Appl	C 533	13.2	66.0	1959	4	US-09-427-770-67	Sequence 67, Appl
C 461	13.2	66.0	1959	3	US-08-993-722A-29	Sequence 29, Appl	C 534	13.2	66.0	1959	4	US-09-427-769-1	Sequence 1, Appl
C 462	13.2	66.0	1959	3	US-08-993-722A-39	Sequence 39, Appl	C 535	13.2	66.0	1959	4	US-09-427-769-3	Sequence 3, Appl
C 463	13.2	66.0	1959	3	US-08-993-722A-41	Sequence 41, Appl	C 536	13.2	66.0	1959	4	US-09-427-769-5	Sequence 5, Appl
C 464	13.2	66.0	1959	3	US-08-993-722A-45	Sequence 45, Appl	C 537	13.2	66.0	1959	4	US-09-427-769-7	Sequence 7, Appl
C 465	13.2	66.0	1959	3	US-08-993-722A-45	Sequence 45, Appl	C 538	13.2	66.0	1959	4	US-09-427-769-9	Sequence 9, Appl

C 539	13.2	66.0	1959	4	US-09-427-769-9	Sequence 9, Appli	612	13.2	66.0	4505	5	PCT-US95-12642-13	Sequence 13, Appli
C 540	13.2	66.0	1959	4	US-09-427-769-11	Sequence 11, Appli	613	13.2	66.0	4526	4	US-09-424-283-7	Sequence 7, Appli
C 541	13.2	66.0	1959	4	US-09-427-769-12	Sequence 12, Appli	614	13.2	66.0	4466	4	US-09-485-286-14	Sequence 14, Appli
C 542	13.2	66.0	1959	4	US-09-427-769-27	Sequence 27, Appli	615	13.2	66.0	4700	2	US-08-928-692-16	Sequence 16, Appli
C 543	13.2	66.0	1959	4	US-09-427-769-29	Sequence 29, Appli	616	13.2	66.0	4700	2	US-08-939-972-16	Sequence 16, Appli
C 544	13.2	66.0	1959	4	US-09-427-769-39	Sequence 39, Appli	617	13.2	66.0	4771	2	US-08-866-650-2	Sequence 2, Appli
C 545	13.2	66.0	1959	4	US-09-427-769-41	Sequence 41, Appli	618	13.2	66.0	4771	2	US-09-021-287-2	Sequence 2, Appli
C 546	13.2	66.0	1959	4	US-09-427-769-43	Sequence 43, Appli	619	13.2	66.0	4771	3	US-09-240-473-2	Sequence 2, Appli
C 547	13.2	66.0	1959	4	US-09-427-769-45	Sequence 45, Appli	620	13.2	66.0	4864	4	US-09-340-798A-1	Sequence 1, Appli
C 548	13.2	66.0	1959	4	US-09-427-769-47	Sequence 47, Appli	621	13.2	66.0	4918	1	US-09-173-053-7	Sequence 7, Appli
C 549	13.2	66.0	1959	4	US-09-427-769-49	Sequence 49, Appli	622	13.2	66.0	4928	1	US-08-345-913-1	Sequence 1, Appli
C 550	13.2	66.0	1959	4	US-09-427-769-53	Sequence 53, Appli	623	13.2	66.0	4928	3	US-08-818-562-1	Sequence 1, Appli
C 551	13.2	66.0	1959	4	US-09-427-769-63	Sequence 63, Appli	624	13.2	66.0	4928	3	US-09-628-445-1	Sequence 1, Appli
C 552	13.2	66.0	1959	4	US-09-427-769-65	Sequence 65, Appli	625	13.2	66.0	4965	2	US-08-564-313-1	Sequence 1, Appli
C 553	13.2	66.0	1959	4	US-09-427-769-67	Sequence 67, Appli	626	13.2	66.0	4965	5	PCT-US94-06069-1	Sequence 1, Appli
C 554	13.2	66.0	1959	4	US-09-427-769-97	Sequence 97, Appli	627	13.2	66.0	5053	4	US-09-311-784A-35	Sequence 35, Appli
C 555	13.2	66.0	1970	3	US-08-687-590-56	Sequence 56, Appli	628	13.2	66.0	5183	3	US-09-039-555B-18	Sequence 18, Appli
C 556	13.2	66.0	1970	6	US-08-687-590-56	Sequence 56, Appli	629	13.2	66.0	5203	4	US-09-976-594-218	Sequence 218, App
C 557	13.2	66.0	1978	4	US-09-525-223A-8	Sequence 8, Appli	630	13.2	66.0	5215	4	US-09-173-053-8	Sequence 8, Appli
C 558	13.2	66.0	1981	4	US-09-647-390-17	Sequence 17, Appli	631	13.2	66.0	5243	2	US-08-414-335-2	Sequence 2, Appli
C 559	13.2	66.0	2091	4	US-09-107-532A-276	Sequence 276, App	632	13.2	66.0	5502	4	US-09-702-705-785	Sequence 785, App
C 560	13.2	66.0	2115	1	US-08-329-681A-3	Sequence 3, Appli	633	13.2	66.0	5502	4	US-09-736-457-785	Sequence 785, App
C 561	13.2	66.0	2280	3	US-08-996-441B-102	Sequence 102, App	634	13.2	66.0	5502	4	US-09-614-124B-785	Sequence 785, App
C 562	13.2	66.0	2280	3	US-08-993-722A-102	Sequence 102, App	635	13.2	66.0	5502	4	US-09-671-325-785	Sequence 785, App
C 563	13.2	66.0	2280	3	US-08-993-170A-102	Sequence 102, App	636	13.2	66.0	5502	4	US-09-589-184-785	Sequence 785, App
C 564	13.2	66.0	2280	3	US-08-993-775B-102	Sequence 102, App	637	13.2	66.0	5592	4	US-09-495-797-17	Sequence 37, Appli
C 565	13.2	66.0	2280	4	US-09-427-770-102	Sequence 102, App	638	13.2	66.0	5676	2	US-08-663-998-3	Sequence 3, Appli
C 566	13.2	66.0	2280	4	US-09-427-769-102	Sequence 102, App	639	13.2	66.0	5682	2	US-08-663-998-4	Sequence 4, Appli
C 567	13.2	66.0	2392	4	US-09-797-464A-1	Sequence 1, Appli	640	13.2	66.0	5838	2	US-08-850-860-3	Sequence 3, Appli
C 568	13.2	66.0	2430	5	PCT-US92-00040-1	Sequence 1, Appli	641	13.2	66.0	5838	2	US-08-944-916-3	Sequence 3, Appli
C 569	13.2	66.0	2479	4	US-09-050-863-4	Sequence 4, Appli	642	13.2	66.0	5838	2	US-08-814-877-3	Sequence 3, Appli
C 570	13.2	66.0	2479	4	US-09-359-081-4	Sequence 4, Appli	643	13.2	66.0	5838	4	US-09-272-432A-3	Sequence 3, Appli
C 571	13.2	66.0	2511	4	US-09-620-312D-559	Sequence 559, App	644	13.2	66.0	5845	4	US-09-173-053-1	Sequence 1, Appli
C 572	13.2	66.0	2622	4	US-09-543-681A-5143	Sequence 1313, Ap	645	13.2	66.0	5899	4	US-09-173-053-2	Sequence 2, Appli
C 573	13.2	66.0	2624	4	US-09-370-838A-19	Sequence 19, Appli	646	13.2	66.0	5900	2	US-08-663-998-1	Sequence 1, Appli
C 574	13.2	66.0	2632	2	US-08-899-324-32	Sequence 32, Appli	647	13.2	66.0	5902	4	US-09-381-989-4	Sequence 4, Appli
C 575	13.2	66.0	2632	2	US-08-832-892B-32	Sequence 32, Appli	648	13.2	66.0	5926	4	US-09-027-169-3	Sequence 3, Appli
C 576	13.2	66.0	2885	4	US-10-164-595-39	Sequence 39, Appli	649	13.2	66.0	5926	4	US-09-027-169-4	Sequence 4, Appli
C 577	13.2	66.0	3224	4	US-08-965-728A-2	Sequence 2, Appli	650	13.2	66.0	5952	2	US-08-663-998-2	Sequence 2, Appli
C 578	13.2	66.0	3311	4	US-10-164-595-11	Sequence 11, Appli	651	13.2	66.0	6044	1	US-08-316-950-18	Sequence 18, Appli
C 579	13.2	66.0	3356	1	US-08-246-403A-1	Sequence 1, Appli	652	13.2	66.0	6044	5	PCT-US95-12642-18	Sequence 12, Appli
C 580	13.2	66.0	3356	1	US-08-246-403A-4	Sequence 4, Appli	653	13.2	66.0	6047	5	US-08-316-950-12	Sequence 12, Appli
C 581	13.2	66.0	3375	4	US-09-309-572-1	Sequence 1, Appli	654	13.2	66.0	6047	5	PCT-US95-12642-12	Sequence 12, Appli
C 582	13.2	66.0	3375	4	US-09-718-096-1	Sequence 1, Appli	655	13.2	66.0	6171	5	US-08-316-950-17	Sequence 17, Appli
C 583	13.2	66.0	3418	2	US-08-944-916-12	Sequence 12, Appli	656	13.2	66.0	6171	5	PCT-US95-12642-17	Sequence 17, Appli
C 584	13.2	66.0	3424	4	US-09-553-690-3	Sequence 3, Appli	657	13.2	66.0	6353	4	US-09-702-705-784	Sequence 784, App
C 585	13.2	66.0	3470	4	US-09-976-594-345	Sequence 345, App	658	13.2	66.0	6353	4	US-09-736-457-784	Sequence 784, App
C 586	13.2	66.0	3547	4	US-09-340-798A-43	Sequence 43, Appli	659	13.2	66.0	6353	4	US-09-614-124B-784	Sequence 784, App
C 587	13.2	66.0	3610	4	US-09-194-949A-1	Sequence 1, Appli	660	13.2	66.0	6353	4	US-09-671-325-784	Sequence 784, App
C 588	13.2	66.0	3870	2	US-08-853-659A-26	Sequence 26, Appli	661	13.2	66.0	6353	4	US-09-589-184-784	Sequence 784, App
C 589	13.2	66.0	3895	4	US-08-961-527-201	Sequence 201, App	662	13.2	66.0	6407	1	US-07-977-630-1	Sequence 1, Appli
C 590	13.2	66.0	3987	4	US-09-082-649B-83	Sequence 83, Appli	663	13.2	66.0	6407	1	US-07-977-630-2	Sequence 2, Appli
C 591	13.2	66.0	3987	4	US-09-082-649B-84	Sequence 84, Appli	664	13.2	66.0	6873	3	US-09-131-028A-8	Sequence 8, Appli
C 592	13.2	66.0	3993	3	US-08-316-950-14	Sequence 14, Appli	665	13.2	66.0	6873	3	US-09-131-028A-8	Sequence 8, Appli
C 593	13.2	66.0	3993	3	PCT-US95-12642-14	Sequence 14, Appli	666	13.2	66.0	6978	4	US-09-872-733A-15	Sequence 15, Appli
C 594	13.2	66.0	4059	2	US-08-564-313-2	Sequence 2, Appli	667	13.2	66.0	7091	3	US-08-975-762-46	Sequence 46, Appli
C 595	13.2	66.0	4059	2	PCT-US94-06069-2	Sequence 2, Appli	668	13.2	66.0	7091	3	US-09-295-028-46	Sequence 46, Appli
C 596	13.2	66.0	4066	3	US-09-298-367B-3	Sequence 3, Appli	669	13.2	66.0	7091	4	US-09-106-582-46	Sequence 46, Appli
C 597	13.2	66.0	4120	1	US-07-977-630-20	Sequence 20, Appli	670	13.2	66.0	7091	4	US-09-159-469-46	Sequence 46, Appli
C 598	13.2	66.0	4120	1	US-07-977-630-21	Sequence 21, Appli	671	13.2	66.0	7091	4	US-09-693-542-46	Sequence 46, Appli
C 599	13.2	66.0	4120	1	US-08-316-950-11	Sequence 11, Appli	672	13.2	66.0	7612	4	US-09-700-934A-72	Sequence 2, Appli
C 600	13.2	66.0	4120	5	PCT-US95-12642-11	Sequence 11, Appli	673	13.2	66.0	7641	1	US-07-924-028A-5	Sequence 5, Appli
C 601	13.2	66.0	4261	4	US-09-483-419-2	Sequence 2, Appli	674	13.2	66.0	7676	4	US-09-036-356-113	Sequence 213, App
C 602	13.2	66.0	4261	4	US-09-194-949A-35	Sequence 25, Appli	675	13.2	66.0	7676	4	US-09-072-556-708	Sequence 213, App
C 603	13.2	66.0	4296	5	PCT-US95-12642-15	Sequence 15, Appli	676	13.2	66.0	7676	4	US-09-287-849-9	Sequence 9, Appli
C 604	13.2	66.0	4326	3	US-08-760-615-7	Sequence 7, Appli	677	13.2	66.0	8031	4	US-09-643-597-254	Sequence 254, App
C 605	13.2	66.0	4326	3	US-08-015-986A-1	Sequence 1, Appli	678	13.2	66.0	8031	4	US-09-480-884A-254	Sequence 254, App
C 606	13.2	66.0	4338	1	US-08-446-363-1	Sequence 1, Appli	679	13.2	66.0	8031	4	US-09-542-615A-254	Sequence 254, App
C 607	13.2	66.0	4352	2	US-08-316-950-16	Sequence 16, Appli	680	13.2	66.0	8031	4	US-09-606-421B-254	Sequence 254, App
C 608	13.2	66.0	4352	5	PCT-US95-12642-16	Sequence 16, Appli	681	13.2	66.0	8366	4	US-09-872-733A-6	Sequence 6, Appli
C 609	13.2	66.0	4427	3	US-09-342-353-2	Sequence 2, Appli	682	13.2	66.0	8430	3	US-09-131-028A-6	Sequence 6, Appli
C 610	13.2	66.0	4505	1	US-08-316-950-13	Sequence 13, Appli	683	13.2	66.0	8430	3	US-09-131-028A-10	Sequence 10, Appli
C 611	13.2	66.0	4505	1	US-08-316-950-13	Sequence 13, Appli	684	13.2	66.0	8430	3	US-09-131-028A-10	Sequence 10, Appli

C 685	13.2	65.0	8501.3	US-08-793-900-1	Sequence 1, App1	758	12.8	64.0	301.4	US-09-702-705-1714	Sequence 1714, App
C 686	13.2	66.0	8867.2	US-08-853-659A-6	Sequence 6, App1	759	12.8	64.0	301.4	US-09-726-457-1714	Sequence 1714, App
C 687	13.2	66.0	8867.2	US-08-853-659A-9	Sequence 9, App1	760	12.8	64.0	301.4	US-09-671-325-1714	Sequence 1714, App
C 688	13.2	66.0	8867.2	US-08-853-659A-67	Sequence 67, App1	761	12.8	64.0	304.4	US-09-437-261-4	Sequence 4, App1
C 689	13.2	66.0	8967.2	US-08-853-659A-67	Sequence 67, App1	762	12.8	64.0	304.4	US-09-427-613-4	Sequence 4, App1
C 690	13.2	66.0	9048.4	US-08-973-273-4	Sequence 4, App1	763	12.8	64.0	354.4	US-09-250-609-100	Sequence 100, App
C 691	13.2	66.0	10254.4	US-08-961-527-29	Sequence 29, App1	764	12.8	64.0	354.4	US-09-250-611-100	Sequence 100, App
C 692	13.2	66.0	10317.3	US-09-058-746-1	Sequence 1, App1	765	12.8	64.0	357.4	US-09-107-552A-93	Sequence 93, App1
C 693	13.2	66.0	10317.3	US-09-438-142-1	Sequence 1, App1	766	12.8	64.0	396.4	US-09-640-173-153	Sequence 153, App
C 694	13.2	66.0	10323.4	US-09-880-428A-11	Sequence 11, App1	767	12.8	64.0	396.4	US-09-713-550-153	Sequence 153, App
C 695	13.2	66.0	12284.2	US-08-876-991-1	Sequence 1, App1	768	12.8	64.0	414.4	US-09-621-976-18437	Sequence 18437, A
C 696	13.2	66.0	12284.2	US-09-059-853-1	Sequence 1, App1	769	12.8	64.0	425.4	US-09-221-017B-3344	Sequence 3344, A
C 697	13.2	66.0	12311.3	US-08-150-717-1	Sequence 1, App1	770	12.8	64.0	425.4	US-09-250-609-101	Sequence 101, App
C 698	13.2	66.0	12665.4	US-08-961-527-134	Sequence 134, App	771	12.8	64.0	425.4	US-09-250-611-101	Sequence 101, App
C 699	13.2	66.0	12785.4	US-09-553-690-1	Sequence 1, App1	772	12.8	64.0	449.4	US-09-713-550-197	Sequence 197, App
C 700	13.2	66.0	14113.3	US-09-423-134-1	Sequence 1, App1	773	12.8	64.0	454.4	US-08-991-789A-211	Sequence 211, App
C 701	13.2	66.0	14113.3	US-08-992-801-1	Sequence 1, App1	774	12.8	64.0	454.4	US-09-062-451-211	Sequence 211, App
C 702	13.2	66.0	14113.3	US-09-423-535-1	Sequence 1, App1	775	12.8	64.0	454.4	US-09-599-326-211	Sequence 211, App
C 703	13.2	66.0	14194.4	US-09-577-424-3	Sequence 3, App1	776	12.8	64.0	454.4	US-09-289-138-211	Sequence 211, App
C 704	13.2	66.0	15538.4	US-09-554-337-1	Sequence 1, App1	777	12.8	64.0	454.4	US-09-429-755-211	Sequence 211, App
C 705	13.2	66.0	22846.2	US-08-469-461-3	Sequence 3, App1	778	12.8	64.0	477.4	US-09-252-991A-5351	Sequence 5351, App
C 706	13.2	66.0	22846.3	US-07-890-609-3	Sequence 3, App1	779	12.8	64.0	487.4	US-09-030-607-206	Sequence 206, App
C 707	13.2	66.0	24417.2	US-08-446-762-1	Sequence 1, App1	780	12.8	64.0	487.4	US-09-439-313-206	Sequence 206, App
C 708	13.2	66.0	24701.2	US-08-853-659A-2	Sequence 2, App1	781	12.8	64.0	487.4	US-09-352-616A-206	Sequence 206, App
C 709	13.2	66.0	24701.2	US-08-853-659A-3	Sequence 3, App1	782	12.8	64.0	487.4	US-09-232-199A-206	Sequence 206, App
C 710	13.2	66.0	24701.2	US-08-853-659A-60	Sequence 60, App1	783	12.8	64.0	487.4	US-09-159-812-206	Sequence 206, App
C 711	13.2	66.0	24701.2	US-08-853-659A-61	Sequence 61, App1	784	12.8	64.0	487.4	US-09-636-215-206	Sequence 206, App
C 712	13.2	66.0	26016.4	US-09-126-480A-1	Sequence 1, App1	785	12.8	64.0	487.4	US-09-685-166A-206	Sequence 206, App
C 713	13.2	66.0	28001.4	US-09-819-293-3	Sequence 3, App1	786	12.8	64.0	487.4	US-09-115-453-206	Sequence 206, App
C 714	13.2	66.0	30101.4	US-10-193-293-3	Sequence 3, App1	787	12.8	64.0	487.4	US-09-688-489-206	Sequence 206, App
C 715	13.2	66.0	30310.4	US-09-657-346A-96	Sequence 96, App1	788	12.8	64.0	495.4	US-09-833-361-1146	Sequence 1146, App
C 716	13.2	66.0	36138.4	US-08-111-731A-136	Sequence 136, App	789	12.8	64.0	501.4	US-09-134-000C-2056	Sequence 2056, App
C 717	13.2	66.0	40085.4	US-08-111-731A-26	Sequence 26, App1	790	12.8	64.0	506.4	US-09-645-168-7	Sequence 7, App1
C 718	13.2	66.0	81001.4	US-09-750-581-10	Sequence 10, App1	791	12.8	64.0	516.4	US-09-250-609-99	Sequence 99, App1
C 719	13.2	66.0	98844.4	US-09-791-211-10	Sequence 10, App1	792	12.8	64.0	516.4	US-09-250-611-99	Sequence 99, App1
C 720	13.2	66.0	100848.4	US-09-596-002-39	Sequence 39, App1	793	12.8	64.0	527.4	US-09-543-661A-2905	Sequence 2905, App
C 721	13.2	66.0	124884.4	US-09-661-596A-76	Sequence 76, App1	794	12.8	64.0	527.4	US-09-621-976-1248	Sequence 1248, App
C 722	13.2	66.0	124884.4	US-09-913-514-1-1	Sequence 1, App1	795	12.8	64.0	540.4	US-09-134-000C-2406	Sequence 2406, App
C 723	13.2	66.0	125157.4	US-09-913-514-2	Sequence 2, App1	796	12.8	64.0	558.4	US-09-328-352-1847	Sequence 1847, App
C 724	13.2	66.0	128779.4	US-09-497-855A-38	Sequence 38, App1	797	12.8	64.0	586.4	US-09-918-656-17	Sequence 17, App1
C 725	13.2	66.0	169998.4	US-09-676-610B-24	Sequence 24, App1	798	12.8	64.0	601.4	US-09-465-559-21	Sequence 21, App1
C 726	13.2	66.0	197496.4	US-09-877-177A-10	Sequence 10, App1	799	12.8	64.0	645.4	US-08-956-111E-686	Sequence 686, App
C 727	13.2	66.0	269223.4	US-09-596-002-41	Sequence 41, App1	800	12.8	64.0	651.4	US-09-328-352-2818	Sequence 2818, App
C 728	13.2	66.0	1664976.4	US-08-916-421B-1	Sequence 1, App1						
C 729	13.2	66.0	422.4	US-09-397-787-221	Sequence 221, App						
C 730	13.2	65.0	474.4	US-09-621-976-19085	Sequence 19085, A						
C 731	13.2	65.0	702.4	US-09-328-352-3803	Sequence 3803, App						
C 732	13.2	65.0	1215.4	US-09-328-352-3884	Sequence 3884, App						
C 733	13.2	65.0	1357.2	US-08-546-117-8	Sequence 8, App1						
C 734	13.2	65.0	2186.4	US-09-620-312D-735	Sequence 735, App						
C 735	13.2	65.0	2308.3	US-09-382-256-9	Sequence 9, App1						
C 736	13.2	65.0	2308.3	US-09-395-115-9	Sequence 9, App1						
C 737	13.2	65.0	2308.4	US-08-436-265-9	Sequence 9, App1						
C 738	13.2	65.0	2308.4	US-09-679-187-9	Sequence 9, App1						
C 739	13.2	65.0	2628.2	US-08-696-944-1	Sequence 1, App1						
C 740	13.2	65.0	2760.4	US-09-543-681A-86	Sequence 86, App1						
C 741	13.2	65.0	3713.1	US-08-330-537-3	Sequence 3, App1						
C 742	13.2	65.0	3713.1	US-08-330-537-4	Sequence 4, App1						
C 743	13.2	65.0	5474.1	US-08-137-252-2	Sequence 2, App1						
C 744	13.2	65.0	6396.4	US-09-620-312D-226	Sequence 226, App						
C 745	13.2	65.0	18385.4	US-08-956-171E-67	Sequence 67, App1						
C 746	13.2	65.0	148567.4	US-09-801-876B-3	Sequence 3, App1						
C 747	13.2	65.0	148567.4	US-10-254-869-3	Sequence 3, App1						
C 748	12.8	64.0	21.4	US-09-422-978-9330	Sequence 9320, App						
C 749	12.8	64.0	201.4	US-09-404-879A-335	Sequence 345, App						
C 750	12.8	64.0	221.4	US-09-252-991A-5326	Sequence 526, App						
C 751	12.8	64.0	258.4	US-09-543-681A-3378	Sequence 3278, App						
C 752	12.8	64.0	259.3	US-09-172-108-35	Sequence 35, App1						
C 753	12.8	64.0	259.3	US-09-172-111-33	Sequence 33, App1						
C 754	12.8	64.0	274.4	US-09-313-294A-1666	Sequence 1666, App						
C 755	12.8	64.0	279.4	US-09-489-039A-6634	Sequence 6634, App						
C 756	12.8	64.0	281.4	US-09-313-294A-1178	Sequence 1178, App						
C 757	12.8	64.0	298.4	US-09-113-294A-4029	Sequence 4029, App						
ALIGNMENTS											
RESULT 1											
US-09-651-656-8/C											
; Sequence 8, Application US/09651656											
; Patent No. 6340566											
; GENERAL INFORMATION:											
; APPLICANT: MCCUTHEN-MALONEY, SANDRA											
; APPLICANT: LAWRENCE DETMORE NATIONAL LABORATORY											
; TITLE OF INVENTION: DETECTION AND QUANTIFICATION OF SINGLE NUCLEOTIDE											
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS											
; FILE REFERENCE: IL-10689											
; CURRENT APPLICATION NUMBER: US/09/651,656											
; PRIORITY FILING DATE: 2000-08-29											
; PRIOR APPLICATION NUMBER: 60/192,764											
; NUMBER OF SEQ ID NOS: 106											
; SOFTWARE: PatentIn Ver. 2.1											
; SEQ ID NO 8											
; LENGTH: 822											
; TYPE: DNA											
; ORGANISM: Homo sapiens											
US-09-651-656-8											
Query Match 100.0%; Score 20; DB 4; Length 822;											
Best Local Similarity 100.0%; Pred. No. 0.48;											

```

RESULT 1
US-09-651-656-8/c
; Sequence 8, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651,656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-651-656-8

Query Match      100.0%; Score 20; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.48;

```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCATCACTCATGTTGATG 20
|||||

Db 743 GGTCATCACTCATGTTGATG 724

RESULT 2
US-09-650-855-8/c
; Sequence 8, Application US/09650855
; Patent No. 636535
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650,855
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-650-855-8

Query Match 100.0%; Score 20; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCATCACTCATGTTGATG 20
|||||

Db 743 GGTCATCACTCATGTTGATG 724

RESULT 3
US-08-961-527-40/c
; Sequence 40, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kueh
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-40

Query Match 84.0%; Score 16.8; DB 4; Length 14273;
Best Local Similarity 90.0%; Pred. No. 31;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCATCACTCATGTTGATG 20
|||||

Db 13385 GGTCATCACTCATGTTGATG 13366

RESULT 4
US-09-385-982-62
; Sequence 62, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(614)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-62

Query Match 79.0%; Score 15.8; DB 3; Length 614;
Best Local Similarity 89.5%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTCCATCACTCATGTTGATG 20
|||||

Db 466 GTCCATCACTCATGTTGATG 484

RESULT 5
US-09-023-655-1225/c
; Sequence 1225, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Selhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/023,655
/ FILING DATE: HEREWITH
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1225:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 686 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g246741
/
/ US-09-023-655-1225
```

```
Query Match          79.0%; Score 15.8; DB 4; Length 686;
Best Local Similarity 89.5%; Pred. No. 61;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCACTCATGTTGAT 19
          |||||
DB      628 GGTCCATCACTCATGATGAT 610
```

```
RESULT 6
US-09-543-681A-137/C
/ Sequence 137, Application US/09543681A
/ Patent No. 6605709
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETTON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/09/543,681A
/ CURRENT FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 137
/ LENGTH: 801
/ TYPE: DNA
/ ORGANISM: Proteus mirabilis
/
/ US-09-543-681A-137
```

```
Query Match          79.0%; Score 15.8; DB 4; Length 801;
Best Local Similarity 89.5%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GTTCATCACTCATGTTGATG 20
          |||||
DB      136 GACCAACTCATGTTGATG 118
```

```
RESULT 7
US-09-543-681A-1773/C
/ Sequence 1773, Application US/09543681A
/ Patent No. 6605709
```

```
/
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETTON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/09/543,681A
/ CURRENT FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 1773
/ LENGTH: 1176
/ TYPE: DNA
/ ORGANISM: Proteus mirabilis
/
/ US-09-543-681A-1773
```

```
Query Match          79.0%; Score 15.8; DB 4; Length 1176;
Best Local Similarity 89.5%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GTTCATCACTCATGTTGATG 20
          |||||
DB      417 GTTCATCACTCATGATGATG 399
```

```
RESULT 8
US-09-134-001C-388
/ Sequence 388, Application US/09134001C
/ Patent No. 6380370
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ FILE REFERENCE: GTC-007
/ CURRENT APPLICATION NUMBER: US/09/134,001C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ PRIOR FILING DATE: 1997-08-14
/ NUMBER OF SEQ ID NOS: 5674
/ SEQ ID NO 388
/ LENGTH: 675
/ TYPE: DNA
/ ORGANISM: Staphylococcus epidermidis
/
/ US-09-134-001C-388
```

```
Query Match          77.0%; Score 15.4; DB 4; Length 675;
Best Local Similarity 94.1%; Pred. No. 96;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 GTTCATCACTCATGTTGA 18
          |||||
DB      427 GTTCATCACTCATGTTGA 443
```

```
RESULT 9
US-08-956-171E-1137
/ Sequence 1137, Application US/08956171E
/ Patent No. 6593114
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunesh
/ GILL H. CHOI
/ PATRICK S. DILLON
/ CRAIG A. ROSEN
/ STEVEN C. BARASH
/ MICHAEL R. FANNON
/
/ TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 5256
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
```

STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 1137:
SEQUENCE CHARACTERISTICS:
LENGTH: 330 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1137:
US-08-956-171E-1137

Query Match 76.0%; Score 15.2; DB 4; Length 330;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGTGATG 20
DB 284 GGACCAACTCATGTGTGATG 303

RESULT 10
US-09-489-039A-5224/C
Sequence 5224, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIORITY FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 5224
LENGTH: 1419
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5224

Query Match 76.0%; Score 15.2; DB 4; Length 1419;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGTGATG 20
DB 487 GGTCATCTCATGTGTGATG 468

RESULT 11
US-09-328-352-2478

Sequence 2478, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
PRIORITY FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 2478
LENGTH: 4884
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-2478

Query Match 76.0%; Score 15.2; DB 4; Length 4884;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGTGATG 20
DB 1128 GTCTTACTCATGTGTGATG 1147

RESULT 12
US-08-646-695-1
Sequence 1, Application US/08646695
Patent No. 6168943
GENERAL INFORMATION:
APPLICANT: Rose, John K.
TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESSES:
ADDRESSER: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,695
FILING DATE: On Even Date Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8664
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1431 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 760..2025
FEATURE:
NAME/KEY: CDS
LOCATION: 2092..2886
FEATURE:
NAME/KEY: CDS

/ LOCATION: 2946..3632
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3774..5306
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 5429..11755
/ US-08-646-695-1

Query Match 76.0%; Score 15.2; DB 3; Length 14311;
Best Local Similarity 85.0%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
DB 11703 GGTCCTACTCATGTTGATG 11722

RESULT 13
US-08-646-695-7/c

/ Sequence 7, Application US/08646695
/ Patent No. 6168943
/ GENERAL INFORMATION:
/ APPLICANT: Rose, John K.
/ TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
/ TITLE OF INVENTION: US95
/ NUMBER OF SEQUENCES: 44
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: PENNIE & EDMONDS
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/646,695
/ FILING DATE: On Even Date Herewith
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mirock, S. Leslie
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 6523-008
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 14311 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ US-08-646-695-7

Query Match 76.0%; Score 15.2; DB 3; Length 14311;
Best Local Similarity 85.0%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
DB 2609 GGTCCTACTCATGTTGATG 2590

RESULT 14
PCT-US96-06053-1

/ Sequence 1, Application PC/TUS9606053
/ GENERAL INFORMATION:

/ APPLICANT: Yale University
/ TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
/ TITLE OF INVENTION: US95
/ NUMBER OF SEQUENCES: 41
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: PENNIE & EDMONDS
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/06053
/ FILING DATE: 01-MAY-1996
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mirock, S. Leslie
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 6523-009-228
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 14311 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA

/ FEATURE: CDS
/ NAME/KEY: 760..2028
/ LOCATION: 2092..2889
/ FEATURE: CDS
/ NAME/KEY: 2946..3635
/ LOCATION: 3774..5309
/ FEATURE: CDS
/ NAME/KEY: 5429..11758
/ LOCATION: 5429..11758
/ PCT-US96-06053-1

Query Match 76.0%; Score 15.2; DB 5; Length 14311;
Best Local Similarity 85.0%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
DB 11703 GGTCCTACTCATGTTGATG 11722

RESULT 15

PCT-US96-06053-7/c

/ Sequence 7, Application PC/TUS9606053
/ GENERAL INFORMATION:
/ APPLICANT: Yale University
/ TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
/ TITLE OF INVENTION: US95
/ NUMBER OF SEQUENCES: 41
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: PENNIE & EDMONDS
/ STREET: 1155 Avenue of the Americas
/ CITY: New York

[illegible]

FILE REFERENCE: GENSET.054PR2

```
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 3741
/ LENGTH: 387
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 191..376
US-09-621-976-3741

Query Match
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TCCATCTCATGTTGATG 20
Db 182 TCCAGCTCATGTTGATG 199

RESULT 23
US-09-328-352-992
/ Sequence 992, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 992
/ LENGTH: 546
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-992

Query Match
Best Local Similarity 88.9%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TCCATCTCATGTTGATG 20
Db 10 TCTATGCTCATGTTGATG 27

RESULT 24
US-09-489-039A-911
/ Sequence 911, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ FILE REFERENCE: 2709.2004001
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 911
/ LENGTH: 2244
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-911

Query Match
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20
```

```
Db 1711 CATGCTCATGTTGATG 1726

RESULT 25
US-09-489-039A-953/C
/ Sequence 953, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ FILE REFERENCE: 2709.2004001
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 953
/ LENGTH: 2310
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-953

Query Match
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20
Db 969 CATGCTCATGTTGATG 954

RESULT 26
US-08-136-743B-40
/ Sequence 40, Application US/08136743B
/ Patent No. 5459063
/ GENERAL INFORMATION:
/ APPLICANT: Barry S. Cooperman, Harvey Rubin,
/ TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
/ TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit
/ NUMBER OF SEQUENCES: 67
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: The University of Pennsylvania
/ STREET: Suite 330
/ STREET: 3700 Market Street
/ CITY: Philadelphia
/ STATE: Pennsylvania
/ COUNTRY: U.S.A.
/ ZIP: 19104-3246
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
/ COMPUTER: IBM PS/2
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/136,743B
/ FILING DATE: 10/14/93
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Monaco, Daniel A.
/ REGISTRATION NUMBER: 30,480
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (215) 568-8383
/ TELEFAX: (215) 568-5549
/ TELEX: No. 5459063e
/ INFORMATION FOR SEQ ID NO: 40:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2433 nucleotides
/ TYPE: nucleic acid
```

STRANDEDNESS: single
TOPOLOGY: linear
US-08-136-743B-40

Query Match 72.0%; Score 14.4; DB 1; Length 2433;
Best Local Similarity 93.8%; Pred. No. 3.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATCTGTG 17
DB 2272 GTCCATCTCATCTGTG 2287

RESULT 27
US-08-136-743B-3
Sequence 3, Application US/08136743B
Patent No. 5459063
GENERAL INFORMATION:
APPLICANT: Barry S. Cooperman, Harvey Rubin,
Jerome Salem, and Allison L. Fisher
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: The University of Pennsylvania
STREET: Suite 330
STREET: 3700 Market Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19104-3246
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,743B
FILING DATE: 10/14/93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 3957-10
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5459063e
INFORMATION FOR SEQ. ID NO. 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2663 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-136-743B-3

Query Match 72.0%; Score 14.4; DB 1; Length 2663;
Best Local Similarity 93.8%; Pred. No. 3.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATCTGTG 17
DB 2499 GTCCATCTCATCTGTG 2514

RESULT 28
US-08-247-901C-1
Sequence 1, Application US/08247901C
Patent No. 5750384
GENERAL INFORMATION:
APPLICANT: Jacobs et al
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,901C
FILING DATE: May 23, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/057,531
FILING DATE: April 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/273
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 50341
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: L5 shuttle phasmid sequence
HYPOTHETICAL: NO
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: L5 mycobacteriophage
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION: No. 5750384e
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-247-901C-1

Query Match 72.0%; Score 14.4; DB 1; Length 50341;
Best Local Similarity 93.8%; Pred. No. 6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORGANISM: mycobacteriophage L5
STRAIN: not applicable
INDIVIDUAL ISOLATE: L5
DEVELOPMENTAL STAGE: not applicable
HAPLOTYPE: not applicable
TISSUE TYPE: not applicable
CELL TYPE: not applicable
ORGANELLE: not applicable
IMMEDIATE SOURCE: mycobacteriophage L5 particles
POSITION IN GENOME: entire genome
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Hatfull and Sarkis
TITLE: DNA Sequence, Structure and Gene
TITLE: Expression of Mycobacteriophage L5:
TITLE: A Phage System for Mycobacterial
TITLE: Genetics
JOURNAL: Molecular Microbiology
VOLUME: 7
PAGES: 395-405
DATE: 1993
US-09-426-436-1

Query Match 72.0%; Score 14.4; DB 3; Length 52297;
Best Local Similarity 93.8%; Pred. No. 6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20
DB 9132 CATCCTCATGTTGATG 9147

RESULT 31

US-08-705-557-1
Sequence 1, Application US/08705557
Patent No. 6300061
GENERAL INFORMATION:
APPLICANT: William R. Jacobs, Jr.
APPLICANT: Barry R. Bloom
APPLICANT: Graham F. Hatfull
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ametier, Rochstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,557
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/057,531
FILING DATE:
APPLICATION NUMBER: 07/833,431
FILING DATE: February 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 96700/238

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 52297
TYPE: nucleotide
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: phage genome sequence
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: not applicable.
ORIGINAL SOURCE:
ORGANISM: mycobacteriophage L5
STRAIN: not applicable
INDIVIDUAL ISOLATE: L5
DEVELOPMENTAL STAGE: not applicable
HAPLOTYPE: not applicable
TISSUE TYPE: not applicable
CELL TYPE: not applicable
CELL LINE: not applicable
ORGANELLE: not applicable
IMMEDIATE SOURCE: mycobacteriophage L5 particles
POSITION IN GENOME: entire genome
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Hatfull and Sarkis
TITLE: DNA Sequence, Structure and Gene
TITLE: Expression of Mycobacteriophage L5:
TITLE: A Phage System for Mycobacterial
TITLE: Genetics
JOURNAL: Molecular Microbiology
VOLUME: 7
PAGES: 395-405
DATE: 1993
US-08-705-557-1

Query Match 72.0%; Score 14.4; DB 4; Length 52297;
Best Local Similarity 93.8%; Pred. No. 6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20
DB 9132 CATCCTCATGTTGATG 9147

RESULT 32
US-09-135-782-10
Sequence 10, Application US/09135782
Patent No. 6027929
GENERAL INFORMATION:
APPLICANT: Xu, Shuang-Yong
TITLE OF INVENTION: Method For Cloning And Producing The NspI Restriction
TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The
TITLE OF INVENTION: Recombinant NspI Restriction Endonuclease
FILE REFERENCE: NEB-143
CURRENT APPLICATION NUMBER: US/09/135,782
CURRENT FILING DATE: 1998-08-18
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 36
TYPE: DNA
ORGANISM: No. 6027929loc sp.
US-09-135-782-10

	Query Match	Similarity	Score	No. DB	Length
	Best Local	84.2%	Pred. No. 2	4e-03	36
Matches	16	Conservative	0	Mismatches	3
				Indels	0
				Gaps	0
Qy	1	GGTCCATCTCATGTTGAT	19		
Db	5	GATCCTTACTATATGTTGAT	23		

```

RESULT 33
US-08-956-171E-986
; Sequence 986, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: P248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1124
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 986:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
;
SEQUENCE DESCRIPTION: SEQ ID NO: 986:
US-08-956-171E-986

Query Match 71.0%; Score 14.2; DB 4; Length 329;
Best Local Similarity 84.2%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGTCCATCTCATGTTGAT 19
|||||
|||
Db 228 GGTCAATCTCATTTATGAT 246

RESULT 34
US-09-134-001C-2448/c
; Sequence 2448, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:

```

```

APPLICANT: ymn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2448
LENGTH: 378
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2448

Query Match      71.0%; Score 14.2; DB 4; Length 378;
Best Local Similarity 84.2%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 GTCCTACTCATGTTGATG 20
|||||
193 GACCATCTCATGATGATG 175

RESULT 35
US-09-252-991A-11719
Sequence 11719, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11719
LENGTH: 423
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11719

Query Match      71.0%; Score 14.2; DB 4; Length 423;
Best Local Similarity 84.2%; Pred. No. 3.6e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 GTCCATCTCATGTTGATG 20
|||||
77 GTCCATCTCATGTTGATG 95

RESULT 36
US-09-252-991A-13151
Sequence 13151, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142

```

```

RESULT 35
US-09-252-991A-11719
; Sequence 11719, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11719
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11719

Query Match      71.0%; Score 14.2; DB 4; Length 423;
Best Local Similarity 84.2%; Pred. No. 3.6e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      2  GTCCATCTCATGTTGGATG 20
      ||||| ||||| |||||
DB      77  GTCCATGCTCATGTCGATG 95

RESULT 36
US-09-252-991A-13151
; Sequence 13151, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
;

```

```

US-08-956-171E-986

Query Match      71.0%; Score 14.2; DB 4; Length 329;
Best Local Similarity 84.2%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GGTCCATCTCAGTGTGAT 19
        ||||| ||||| |||||
Db       228 GGTCAATACCTATTATGAT 246

RESULT 34
US-09-134-001C-2448/C
; Sequence 2448, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:

US-09-252-991A-13151
; Sequence 13151, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```



```
; SEQ ID NO 13151
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13151

Query Match
Best Local Similarity 71.0%; Score 14.2; DB 4; Length 483;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCCATCATGTTGAT 19
Db 26 GATCCATGCTCATGTTGAT 44

RESULT 37
US-09-134-000C-244
; Sequence 244, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 244
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (330)..(491)
; OTHER INFORMATION: Nucleotides 330, 382, 430, 440, 449 & 491 are "n" wherein
; OTHER INFORMATION: "n" = any nucleotide.
; OTHER INFORMATION: 1
US-09-134-000C-244

Query Match
Best Local Similarity 71.0%; Score 14.2; DB 4; Length 510;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GTCCTACTCATGTTGAT 20
Db 187 GTCGATCATCATGTTGAT 205

RESULT 38
US-09-107-532A-735
; Sequence 735, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 735:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 555 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...555
; SEQUENCE DESCRIPTION: SEQ ID NO: 735:
US-09-107-532A-735

Query Match
Best Local Similarity 71.0%; Score 14.2; DB 4; Length 555;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCCATCATGTTGAT 19
Db 189 GTCATCCGATGTTGAT 207

RESULT 39
US-09-342-681C-11/c
; Sequence 11, Application US/09342681C
; Patent No. 6355782
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrated ectodermal dysplasia genes and proteins
; FILE REFERENCE: 52978
; CURRENT APPLICATION NUMBER: US/09/342,681C
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (302)
; OTHER INFORMATION: n represents a, c, t, or g
US-09-342-681C-11

Query Match
Best Local Similarity 71.0%; Score 14.2; DB 4; Length 743;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GTCCTACTCATGTTGAT 20
Db 538 GTCCTGCTCATGTTGAT 520
```

RESULT 40
US-09-252-991A-9153/c
; Sequence 9153, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9153
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9153

Query Match 71.0%; Score 14.2; DB 4; Length 909;
Best Local Similarity 84.2%; Pred. No. 4.1e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GTCCATCTCATGTTGATG 20
DB 265 GTCCATGCCCATGCTGATG 247

RESULT 41
US-09-634-238-121
; Sequence 121, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.10431
; CURRENT APPLICATION NUMBER: US/09/634,238
; PRIOR FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Lactobacillus rhamnosus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1065)
; OTHER INFORMATION: n = A,T,C or G
US-09-634-238-121

Query Match 71.0%; Score 14.2; DB 4; Length 1065;
Best Local Similarity 84.2%; Pred. No. 4.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGTCATCTCATGTTGAT 19
DB 224 GGTCATCTCATGCGGAT 242

RESULT 42
US-09-342-681C-14/c
; Sequence 14, Application US/09342681C
; Patent No. 6355782
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydratic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 52978
; CURRENT APPLICATION NUMBER: US/09/342,681C
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1176)
US-09-342-681C-14

Query Match 71.0%; Score 14.2; DB 4; Length 1176;
Best Local Similarity 84.2%; Pred. No. 4.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GTCCATCTCATGTTGATG 20
DB 1128 GTGCTTGTCTCATGTTGATG 1110

RESULT 43
US-09-252-991A-11577/c
; Sequence 11577, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11577
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11577

Query Match 71.0%; Score 14.2; DB 4; Length 1185;
Best Local Similarity 84.2%; Pred. No. 4.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GTCCATCTCATGTTGATG 20
DB 490 GTCCATCTCATGTTGATG 472

RESULT 44
US-08-286-020-1
; Sequence 1, Application US/08286020
; Patent No. 5539095
; GENERAL INFORMATION:
; APPLICANT: Masomeh B. Sticklen and Ravindra K. Hajeja
; TITLE OF INVENTION: A Chitinase cDNA Clone From a

;; TITLE OF INVENTION: Disease Resistant American Elm Tree
;; NUMBER OF SEQUENCES: 1
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ian C. McLeod
;; STREET: 2190 Commons Parkway
;; CITY: Okemos
;; STATE: Michigan
;; COUNTRY: USA
;; ZIP: 48864
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
;; COMPUTER: Acer
;; OPERATING SYSTEM: MS-DOS 5.00
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/286,020
;;
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;;
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ian C. McLeod
;; REGISTRATION NUMBER: 20,931
;; REFERENCE/DOCKET NUMBER: MSU 4.1-207
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (517) 347-4100
;; TELEFAX: (517) 347-4103
;; TELEX: No. 55390956
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1225
;; TYPE: Nucleic Acid
;; STRANDEDNESS: Single
;; TOPOLOGY: Linear
;; MOLECULE TYPE: DNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE:
;; ORIGINAL SOURCE:
;; ORGANISM: Ulmus Americana
;; STRAIN: NPS 3-487
;; INDIVIDUAL ISOLATE: N/A
;; DEVELOPMENTAL STAGE: N/A
;; HAPLOTYPE: N/A
;; TISSUE TYPE: N/A
;; CELL TYPE: N/A
;; CELL LINE: N/A
;; ORGANELLE: N/A
;; IMMEDIATE SOURCE: N/A
;; POSITION IN GENOME: N/A
;; FEATURE:
;; NAME/KEY: chitinase encoding DNA
;; LOCATION:
;; IDENTIFICATION METHOD: sequencing
;; OTHER INFORMATION: DNA needed for chitinase
;; OTHER INFORMATION: in elm.
;; PUBLICATION INFORMATION:
;;
US-08-286-020-1
;
Query Match 71.0%; Score 14.2; DB 1; Length 1225;
Best Local Similarity 84.2%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCATACATCATGTGAT 19
|||||
Db 517 GGTCATACATCTTTGGGAT 535

RESULT 45
US-08-603-919-1
; Sequence 1, Application US/08603919

;; Patent No. 5728382
;; GENERAL INFORMATION:
;; APPLICANT: Masomeh B. Sticklen and Ravindra K. Hajela
;; TITLE OF INVENTION: A Chitinase cDNA Clone From a
;; TITLE OF INVENTION: Disease Resistant American Elm Tree
;; NUMBER OF SEQUENCES: 1
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ian C. McLeod
;; STREET: 2190 Commons Parkway
;; CITY: Okemos
;; STATE: Michigan
;; COUNTRY: USA
;; ZIP: 48864
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
;; COMPUTER: Acer
;; OPERATING SYSTEM: MS-DOS 5.00
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/603,919
;;
;; FILING DATE:
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;;
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ian C. McLeod
;; REGISTRATION NUMBER: 20,931
;; REFERENCE/DOCKET NUMBER: MSU 4.1-207
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (517) 347-4100
;; TELEFAX: (517) 347-4103
;; TELEX: No. 5728382e
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1225
;; TYPE: Nucleic Acid
;; STRANDEDNESS: Single
;; TOPOLOGY: Linear
;; MOLECULE TYPE: DNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE:
;; ORIGINAL SOURCE:
;; ORGANISM: Ulmus Americana
;; STRAIN: NPS 3-487
;; INDIVIDUAL ISOLATE: N/A
;; DEVELOPMENTAL STAGE: N/A
;; HAPLOTYPE: N/A
;; TISSUE TYPE: N/A
;; CELL TYPE: N/A
;; CELL LINE: N/A
;; ORGANELLE: N/A
;; IMMEDIATE SOURCE: N/A
;; POSITION IN GENOME: N/A
;; FEATURE:
;; NAME/KEY: chitinase encoding DNA
;; LOCATION:
;; IDENTIFICATION METHOD: sequencing
;; OTHER INFORMATION: DNA needed for chitinase
;; OTHER INFORMATION: in elm.
;; PUBLICATION INFORMATION:
;;
US-08-603-919-1
;
Query Match 71.0%; Score 14.2; DB 1; Length 1225;
Best Local Similarity 84.2%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCATACATCATGTGAT 19
|||||
Db 517 GGTCATACATCTTTGGGAT 535

RESULT 46
US-08-343-101A-8/c
; Sequence 8, Application US/08343101A
; Patent No. 5810759
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
; TITLE OF INVENTION: Virus Sequences And Uses Thereof
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,101A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-08-343-101A-8

Query Match 71.0%; Score 14.2; DB 2; Length 1315;
Best Local Similarity 84.2%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GTCCATCTCATGTTGATG 20
|||||
Db 805 GTCCCATCTCATGTTGATG 787

RESULT 47
US-09-183-688-8/c
; Sequence 8, Application US/09183688
; Patent No. 6033550
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
; TITLE OF INVENTION: Virus Sequences And Uses Thereof
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/183,688
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/343,101
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; US-09-183-688-8

Query Match 71.0%; Score 14.2; DB 3; Length 1315;
Best Local Similarity 84.2%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GTCCATCTCATGTTGATG 20
|||||
Db 805 GTCCCATCTCATGTTGATG 787

RESULT 48
US-09-519-489-8/c
; Sequence 8, Application US/09519489
; Patent No. 650663
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
; TITLE OF INVENTION: Virus Sequences And Uses Thereof
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/519,489
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,101
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526

;/ INFORMATION FOR SEQ ID NO: 8:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1315 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: DNA (genomic)
;/ HYPOTHETICAL: N
;/ ANTI-SENSE: N
US-09-519-489-8

Query Match 71.0%; Score 14.2; DB 4; Length 1315;
Best Local Similarity 84.2%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
DB 805 GTCCCACTCATGTTGATG 787

RESULT 49
US-09-252-991A-11863
;/ Sequence 11863, Application US/09252991A
;/ Patent No. 6551795
;/ GENERAL INFORMATION:
;/ APPLICANT: Marc J. Rubenfield et al.
;/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;/ FILE REFERENCE: 107196.136
;/ CURRENT APPLICATION NUMBER: US/09/252,991A
;/ CURRENT FILING DATE: 1999-02-18
;/ PRIOR APPLICATION NUMBER: US 60/074,788
;/ PRIOR FILING DATE: 1998-02-18
;/ PRIOR APPLICATION NUMBER: US 60/094,190
;/ PRIOR FILING DATE: 1998-07-27
;/ NUMBER OF SEQ ID NOS: 33142
;/ SEQ ID NO 11863
;/ LENGTH: 1386
;/ TYPE: DNA
;/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11863

Query Match 71.0%; Score 14.2; DB 4; Length 1386;
Best Local Similarity 84.2%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
DB 1248 GTCCATCTCATGTTGATG 1266

RESULT 50
US-09-489-039A-3739
;/ Sequence 3739, Application US/09489039A
;/ Patent No. 6610836
;/ GENERAL INFORMATION:
;/ APPLICANT: Gary Breton et. al
;/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;/ FILE REFERENCE: 2709.2004001
;/ CURRENT APPLICATION NUMBER: US/09/489,039A
;/ CURRENT FILING DATE: 2000-01-27
;/ PRIOR APPLICATION NUMBER: US 60/117,747
;/ PRIOR FILING DATE: 1999-01-29
;/ NUMBER OF SEQ ID NOS: 14342
;/ SEQ ID NO 3739
;/ LENGTH: 1398
;/ TYPE: DNA
;/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3739

Query Match 71.0%; Score 14.2; DB 4; Length 1398;
Best Local Similarity 84.2%; Pred. No. 4.4e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GTCCATCTCATGTTGATG 20
DB 433 GTCCCACTCATGTTGATG 451

Search completed: August 17, 2004, 15:46:45
Job time : 126 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 15:02:58 ; Search time 362 Seconds
(without alignments)
271.083 Million cell updates/sec

Title: US-09-825-489-3

Perfect score: 20

Sequence: 1 ggtccatcacatggtgatg 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 800 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	11	US-09-825-489-3
2	20	100.0	486	9	US-09-867-701-7544
3	20	100.0	506	9	US-09-867-701-2117
4	20	100.0	513	9	US-09-867-701-2107
5	20	100.0	1377	11	US-09-825-489-13
6	20	100.0	1407	15	US-10-103-313-116
7	20	100.0	4670	15	US-10-101-510-587
8	17	85.0	1526	13	US-10-027-632-257254
9	17	85.0	1526	16	US-10-027-632-257254
10	16.8	84.0	14373	13	US-10-158-844-40
11	16.4	82.0	1049	13	US-10-027-632-253417
12	16.4	82.0	1049	16	US-10-027-632-253417
13	15.8	79.0	238	13	US-10-085-783A-40734
14	15.8	79.0	238	16	US-10-242-535A-40734

C 15	15.8	79.0	296	12	US-09-732-627A-2949	Sequence 2949, App
C 16	15.8	79.0	431	15	US-10-060-036-3118	Sequence 3118, App
C 17	15.8	79.0	462	10	US-09-918-995-12216	Sequence 12216, A
C 18	15.8	79.0	540	13	US-10-085-783A-45577	Sequence 45577, A
C 19	15.8	79.0	540	16	US-10-242-535A-45577	Sequence 45577, A
C 20	15.8	79.0	614	10	US-09-871-161-62	Sequence 62, App1
C 21	15.8	79.0	686	17	US-10-641-643-1225	Sequence 1225, App
C 22	15.8	79.0	765	13	US-10-424-599-20394	Sequence 20394, A
C 23	15.8	79.0	789	17	US-10-437-663-64475	Sequence 64475, A
C 24	15.8	79.0	1329	17	US-10-437-663-7531	Sequence 7531, App
C 25	15.8	79.0	3021	12	US-10-041-018-111	Sequence 111, App
C 26	15.8	79.0	5874	12	US-10-041-018-112	Sequence 112, App
C 27	15.8	79.0	101209	13	US-10-087-192-460	Sequence 460, App
C 28	15.4	77.0	519	13	US-10-027-632-79848	Sequence 79848, A
C 29	15.4	77.0	519	13	US-10-027-632-79849	Sequence 79849, A
C 30	15.4	77.0	519	16	US-10-027-632-79848	Sequence 79848, A
C 31	15.4	77.0	519	16	US-10-027-632-79849	Sequence 79849, A
C 32	15.4	77.0	551	9	US-09-764-847-497	Sequence 497, App
C 33	15.4	77.0	551	15	US-10-092-154-497	Sequence 497, App
C 34	15.4	77.0	573	13	US-10-027-632-66707	Sequence 66707, A
C 35	15.4	77.0	573	13	US-10-027-632-66708	Sequence 66708, A
C 36	15.4	77.0	573	13	US-10-027-632-311196	Sequence 311196, A
C 37	15.4	77.0	573	13	US-10-027-632-311197	Sequence 311197, A
C 38	15.4	77.0	573	16	US-10-027-632-66707	Sequence 66707, A
C 39	15.4	77.0	573	16	US-10-027-632-66708	Sequence 66708, A
C 40	15.4	77.0	573	16	US-10-027-632-311196	Sequence 311196, A
C 41	15.4	77.0	573	16	US-10-027-632-311197	Sequence 311197, A
C 42	15.4	77.0	715	13	US-10-424-599-103050	Sequence 103050, A
C 43	15.4	77.0	3767	13	US-10-424-599-23361	Sequence 23361, A
C 44	15.4	77.0	119414	13	US-10-087-192-1807	Sequence 1807, App
C 45	15.4	77.0	335913	10	US-09-754-853A-2	Sequence 2, App11
C 46	15.4	77.0	335913	10	US-09-754-853A-3	Sequence 3, App11
C 47	15.2	76.0	292	13	US-10-424-599-127062	Sequence 127062, A
C 48	15.2	76.0	330	8	US-08-781-986A-1137	Sequence 1137, App
C 49	15.2	76.0	330	13	US-10-329-624-1137	Sequence 1137, App
C 50	15.2	76.0	358	13	US-10-424-599-130364	Sequence 130364, A
C 51	15.2	76.0	449	17	US-10-437-663-31302	Sequence 31302, A
C 52	15.2	76.0	464	13	US-10-424-599-90875	Sequence 90875, A
C 53	15.2	76.0	469	13	US-10-424-599-92079	Sequence 92079, A
C 54	15.2	76.0	572	13	US-10-027-632-208409	Sequence 208409, A
C 55	15.2	76.0	572	16	US-10-027-632-208409	Sequence 208409, A
C 56	15.2	76.0	606	17	US-10-437-663-55289	Sequence 55289, A
C 57	15.2	76.0	630	13	US-10-424-599-128976	Sequence 128976, A
C 58	15.2	76.0	1035	16	US-10-312-773-218	Sequence 218, App
C 59	15.2	76.0	1107	13	US-10-282-122A-26487	Sequence 26487, A
C 60	15.2	76.0	1180	13	US-10-424-599-19603	Sequence 19603, A
C 61	15.2	76.0	1287	9	US-09-738-626-2128	Sequence 2128, App
C 62	15.2	76.0	1335	13	US-10-627-476-621	Sequence 621, App
C 63	15.2	76.0	1410	13	US-10-627-476-619	Sequence 619, App
C 64	15.2	76.0	1513	13	US-10-027-632-253949	Sequence 253949, A
C 65	15.2	76.0	1513	16	US-10-027-632-253949	Sequence 253949, A
C 66	15.2	76.0	1908	15	US-10-156-761-5276	Sequence 5276, App
C 67	15.2	76.0	2000	11	US-09-938-842A-4942	Sequence 4942, App
C 68	15.2	76.0	2000	9	US-09-938-842A-4942	Sequence 4942, App
C 69	15.2	76.0	2913	17	US-10-437-663-17890	Sequence 17890, A
C 70	15.2	76.0	2916	12	US-10-052-482-117	Sequence 117, App
C 71	15.2	76.0	2916	17	US-10-394-948-27	Sequence 27, App1
C 72	15.2	76.0	3334	10	US-09-822-946-102	Sequence 102, App
C 73	15.2	76.0	3670	15	US-10-205-823-125	Sequence 125, App
C 74	15.2	76.0	3717	17	US-10-437-663-91561	Sequence 91561, A
C 75	15.2	76.0	3892	12	US-10-052-482-116	Sequence 116, App
C 76	15.2	76.0	3892	17	US-10-394-948-26	Sequence 26, App1
C 77	15.2	76.0	6222	12	US-09-997-722-24	Sequence 24, App1
C 78	15.2	76.0	6537	15	US-10-175-523-181	Sequence 181, App
C 79	15.2	76.0	6537	15	US-10-175-523-181	Sequence 181, App
C 80	15.2	76.0	7254	12	US-10-267-502-151	Sequence 151, App
C 81	15.2	76.0	17177	15	US-10-225-810-7	Sequence 7, App11
C 82	15.2	76.0	36605	13	US-10-087-192-304	Sequence 304, App
C 83	15.2	76.0	82993	15	US-10-080-170-645	Sequence 645, App
C 84	15.2	76.0	82993	17	US-10-080-170-645	Sequence 645, App
C 85	15.2	76.0	87311	13	US-10-087-192-1417	Sequence 1417, App
C 86	15.2	76.0	96598	12	US-10-052-482-115	Sequence 115, App
C 87	15.2	76.0	96598	17	US-10-394-948-25	Sequence 25, App1

C 88	15.2	76.0	326014	9	US-09-731-231A-3	Sequence 3, Appl1	C 161	14.8	74.0	826	16	US-10-027-632-157166	Sequence 157166,
C 89	15.2	76.0	326014	17	US-10-751-985-3	Sequence 3, Appl1	C 162	14.8	74.0	826	16	US-10-027-632-157167	Sequence 157167,
C 90	15.2	76.0	1230025	16	US-10-289-762-1	Sequence 1, Appl1	C 163	14.8	74.0	835	16	US-10-027-632-8029	Sequence 8029, Ap
C 91	15.2	76.0	3309400	9	US-09-738-626-1	Sequence 1, Appl1	C 164	14.8	74.0	835	16	US-10-027-632-8029	Sequence 8029, Ap
C 92	15.2	76.0	9025608	15	US-10-156-761-1	Sequence 1, Appl1	C 165	14.8	74.0	858	13	US-10-343-650A-533	Sequence 529, App
C 93	15	75.0	331	13	US-10-424-559-961	Sequence 961, App	C 166	14.8	74.0	858	13	US-10-343-650A-533	Sequence 533, App
C 94	15	75.0	593	9	US-09-864-761-7781	Sequence 7781, Ap	C 167	14.8	74.0	868	9	US-10-424-559-54628	Sequence 54628, A
C 95	15	75.0	1851	10	US-09-898-200-10	Sequence 10, Appl1	C 168	14.8	74.0	957	9	US-09-886-055-774	Sequence 374, App
C 96	15	75.0	2000	13	US-09-898-200-2	Sequence 2, Appl1	C 169	14.8	74.0	957	9	US-09-886-055-776	Sequence 376, App
C 97	14.8	74.0	138	13	US-10-282-122A-2025	Sequence 2025, Ap	C 170	14.8	74.0	957	10	US-09-804-291-376	Sequence 376, App
C 98	14.8	74.0	380	9	US-09-864-761-15648	Sequence 15648, A	C 171	14.8	74.0	957	10	US-10-437-863-100467	Sequence 100467,
C 99	14.8	74.0	380	9	US-09-864-761-32154	Sequence 32154, A	C 172	14.8	74.0	1060	17	US-10-437-863-100806	Sequence 100806,
C 100	14.8	74.0	380	9	US-10-282-122A-6873	Sequence 6873, Ap	C 173	14.8	74.0	1075	17	US-10-437-863-100806	Sequence 12041, A
C 101	14.8	74.0	414	13	US-10-424-559-133049	Sequence 133049, A	C 174	14.8	74.0	1181	13	US-10-425-114-12041	Sequence 12041, A
C 102	14.8	74.0	463	9	US-10-424-559-133049	Sequence 133049, A	C 175	14.8	74.0	1282	15	US-10-017-161-597	Sequence 597, App
C 103	14.8	74.0	499	13	US-10-343-887-2262	Sequence 2262, Ap	C 176	14.8	74.0	1258	15	US-10-017-161-597	Sequence 599, App
C 104	14.8	74.0	499	13	US-10-172-118-2262	Sequence 2262, Ap	C 177	14.8	74.0	1258	15	US-10-017-161-597	Sequence 214985,
C 105	14.8	74.0	503	13	US-10-027-632-37952	Sequence 37952, A	C 178	14.8	74.0	1308	13	US-10-027-632-214985	Sequence 523, App
C 106	14.8	74.0	503	13	US-10-027-632-37952	Sequence 37953, A	C 179	14.8	74.0	1308	16	US-10-027-632-214985	Sequence 525, App
C 107	14.8	74.0	503	13	US-10-027-632-65145	Sequence 65145, A	C 180	14.8	74.0	1357	16	US-10-292-798-523	Sequence 135,
C 108	14.8	74.0	503	13	US-10-027-632-65146	Sequence 65146, A	C 181	14.8	74.0	1357	16	US-10-292-798-525	Sequence 42, Appl1
C 109	14.8	74.0	503	13	US-10-027-632-310829	Sequence 310829, A	C 182	14.8	74.0	1364	9	US-09-823-890A-135	Sequence 42, Appl1
C 110	14.8	74.0	503	13	US-10-027-632-310830	Sequence 310830, A	C 183	14.8	74.0	1568	9	US-09-925-299-42	Sequence 42, Appl1
C 111	14.8	74.0	503	16	US-10-027-632-37952	Sequence 37952, A	C 184	14.8	74.0	1568	10	US-09-925-299-42	Sequence 42, Appl1
C 112	14.8	74.0	503	16	US-10-027-632-37952	Sequence 37953, A	C 185	14.8	74.0	1654	13	US-10-425-114-31792	Sequence 40474, A
C 113	14.8	74.0	503	16	US-10-027-632-65145	Sequence 65145, A	C 186	14.8	74.0	1719	13	US-10-282-122A-40474	Sequence 18936, A
C 114	14.8	74.0	503	16	US-10-027-632-65146	Sequence 65146, A	C 187	14.8	74.0	1818	13	US-10-282-122A-18936	Sequence 16939, A
C 115	14.8	74.0	503	16	US-10-027-632-310829	Sequence 310829, A	C 188	14.8	74.0	1836	9	US-09-864-761-16939	Sequence 40328, A
C 116	14.8	74.0	503	16	US-10-027-632-310830	Sequence 310830, A	C 189	14.8	74.0	1938	17	US-10-437-863-40328	Sequence 46591, A
C 117	14.8	74.0	504	17	US-10-021-332-2076	Sequence 2076, Ap	C 190	14.8	74.0	1978	17	US-10-437-863-46591	Sequence 3781, Ap
C 118	14.8	74.0	515	13	US-10-027-632-271233	Sequence 271233, A	C 191	14.8	74.0	1987	9	US-09-864-761-3781	Sequence 1804, Ap
C 119	14.8	74.0	515	16	US-10-027-632-271233	Sequence 271233, A	C 192	14.8	74.0	2000	16	US-10-260-338-1804	Sequence 2354, Ap
C 120	14.8	74.0	528	13	US-10-027-632-50695	Sequence 50695, A	C 193	14.8	74.0	2000	16	US-10-260-338-2354	Sequence 289, App
C 121	14.8	74.0	528	16	US-10-027-632-50695	Sequence 50695, A	C 194	14.8	74.0	2002	9	US-09-887-576-298	Sequence 29, Appl1
C 122	14.8	74.0	536	13	US-10-027-632-85854	Sequence 85854, A	C 195	14.8	74.0	2004	9	US-09-887-576-299	Sequence 54623, A
C 123	14.8	74.0	536	16	US-10-027-632-85854	Sequence 85854, A	C 196	14.8	74.0	2033	17	US-10-466-720-24	Sequence 24, Appl1
C 124	14.8	74.0	556	17	US-10-437-963-68799	Sequence 68799, A	C 197	14.8	74.0	2129	13	US-10-424-599-54625	Sequence 54623, A
C 125	14.8	74.0	564	15	US-10-029-386-4237	Sequence 4237, Ap	C 198	14.8	74.0	2379	11	US-09-938-842A-2171	Sequence 2171, Ap
C 126	14.8	74.0	577	9	US-09-864-761-20547	Sequence 20547, A	C 199	14.8	74.0	2539	11	US-09-938-842A-2171	Sequence 2171, Ap
C 127	14.8	74.0	594	13	US-10-425-114-12543	Sequence 12543, A	C 200	14.8	74.0	2850	13	US-10-027-632-111951	Sequence 111951,
C 128	14.8	74.0	596	13	US-10-027-632-75460	Sequence 75460, A	C 201	14.8	74.0	3061	13	US-10-425-114-6150	Sequence 6150, Ap
C 129	14.8	74.0	596	16	US-10-027-632-75460	Sequence 75460, A	C 202	14.8	74.0	3134	13	US-10-424-599-45533	Sequence 45533, A
C 130	14.8	74.0	600	13	US-10-027-632-208951	Sequence 208951, A	C 203	14.8	74.0	3256	17	US-10-437-863-25305	Sequence 52305, A
C 131	14.8	74.0	600	13	US-10-027-632-208951	Sequence 208951, A	C 204	14.8	74.0	3273	16	US-10-369-493-26879	Sequence 26879, A
C 132	14.8	74.0	600	16	US-10-027-632-208951	Sequence 208951, A	C 205	14.8	74.0	3295	16	US-10-320-797-125	Sequence 125, App
C 133	14.8	74.0	610	13	US-10-027-632-246341	Sequence 246341, A	C 206	14.8	74.0	3545	15	US-10-101-644A-951	Sequence 951, App
C 134	14.8	74.0	610	13	US-10-027-632-246341	Sequence 246341, A	C 207	14.8	74.0	3765	17	US-10-437-863-61646	Sequence 61646, A
C 135	14.8	74.0	610	16	US-10-027-632-246342	Sequence 246342, A	C 208	14.8	74.0	3909	17	US-10-437-863-61646	Sequence 26395, A
C 136	14.8	74.0	610	16	US-10-027-632-246342	Sequence 246342, A	C 209	14.8	74.0	4564	17	US-10-437-863-45318	Sequence 45318, A
C 137	14.8	74.0	610	16	US-10-027-632-246342	Sequence 246342, A	C 210	14.8	74.0	4564	17	US-10-437-863-45318	Sequence 1531, Ap
C 138	14.8	74.0	626	13	US-10-027-632-86952	Sequence 86952, A	C 211	14.8	74.0	4595	9	US-09-917-800A-1531	Sequence 9854, Ap
C 139	14.8	74.0	626	16	US-10-027-632-86952	Sequence 86952, A	C 212	14.8	74.0	5085	15	US-10-198-464-9854	Sequence 9854, Ap
C 140	14.8	74.0	630	13	US-10-027-632-89394	Sequence 89394, A	C 213	14.8	74.0	5110	15	US-10-101-644A-963	Sequence 963, App
C 141	14.8	74.0	630	13	US-10-027-632-89394	Sequence 89394, A	C 214	14.8	74.0	5119	17	US-10-437-863-80960	Sequence 80960, Ap
C 142	14.8	74.0	630	13	US-10-027-632-303989	Sequence 303989, A	C 215	14.8	74.0	6106	16	US-10-062-674-16048	Sequence 16048, Ap
C 143	14.8	74.0	630	13	US-10-027-632-303989	Sequence 303989, A	C 216	14.8	74.0	7566	10	US-09-764-494-10010	Sequence 10010, A
C 144	14.8	74.0	630	16	US-10-027-632-89394	Sequence 89394, A	C 217	14.8	74.0	7741	13	US-10-092-000A-75	Sequence 75, Appl1
C 145	14.8	74.0	630	16	US-10-027-632-89395	Sequence 89395, A	C 218	14.8	74.0	7983	15	US-10-017-161-1987	Sequence 1987, Ap
C 146	14.8	74.0	630	16	US-10-027-632-303989	Sequence 303989, A	C 219	14.8	74.0	29883	16	US-10-292-798-1635	Sequence 1635, Ap
C 147	14.8	74.0	630	16	US-10-027-632-303990	Sequence 303990, A	C 220	14.8	74.0	32757	17	US-10-332-881-140	Sequence 140, App
C 148	14.8	74.0	630	13	US-10-027-632-313264	Sequence 313264, A	C 221	14.8	74.0	32757	17	US-09-921-892-2	Sequence 2, Appl1
C 149	14.8	74.0	636	16	US-10-027-632-313264	Sequence 313264, A	C 222	14.8	74.0	52216	9	US-09-747-810-1	Sequence 1, Appl1
C 150	14.8	74.0	642	13	US-10-282-122A-19342	Sequence 19342, A	C 223	14.8	74.0	90468	16	US-10-085-117-160	Sequence 160, App
C 151	14.8	74.0	656	13	US-10-027-632-279251	Sequence 279251, A	C 224	14.8	74.0	94917	13	US-10-087-192-184	Sequence 184, App
C 152	14.8	74.0	656	16	US-10-027-632-279251	Sequence 279251, A	C 225	14.8	74.0	105319	13	US-10-087-192-658	Sequence 658, App
C 153	14.8	74.0	663	13	US-10-027-632-219251	Sequence 219251, A	C 226	14.8	74.0	126512	9	US-09-804-474A-3	Sequence 3, Appl1
C 154	14.8	74.0	663	16	US-10-027-632-211220	Sequence 211220, A	C 227	14.8	74.0	135827	17	US-10-332-881-1332	Sequence 232, App
C 155	14.8	74.0	728	13	US-10-027-632-142283	Sequence 142283, A	C 228	14.8	74.0	137049	13	US-10-087-192-11198	Sequence 1198, App
C 156	14.8	74.0	728	13	US-10-027-632-142284	Sequence 142284, A	C 229	14.8	74.0	173808	14	US-10-003-806-10	Sequence 53, Appl1
C 157	14.8	74.0	728	16	US-10-027-632-142283	Sequence 142283, A	C 230	14.8	74.0	178896	17	US-10-450-826-53	Sequence 1036, Ap
C 158	14.8	74.0	728	16	US-10-027-632-142284	Sequence 142284, A	C 231	14.8	74.0	296405	13	US-10-087-192-1036	Sequence 1036, Ap
C 159	14.8	74.0	826	13	US-10-027-632-157166	Sequence 157166, A	C 232	14.8	74.0	344548	13	US-10-087-192-334	Sequence 334, App
C 160	14.8	74.0	826	13	US-10-027-632-157167	Sequence 157167, A	C 233	14.8	74.0	418550	16	US-10-292-798-1463	Sequence 1463, App

C 234	14.8	74.0	684973	9	US-09-263-959-1	Sequence 1, Appl1	C 307	14.2	71.0	171	15	US-10-029-386-27336	Sequence 27336, A
C 235	14.8	74.0	1691139	15	US-10-067-514-1	Sequence 1, Appl1	C 308	14.2	71.0	174	13	US-10-085-783A-35599	Sequence 35599, A
C 236	14.8	74.0	1691139	16	US-10-419-723-1	Sequence 1, Appl1	C 309	14.2	71.0	174	16	US-10-242-535A-35599	Sequence 35599, A
C 237	14.4	72.0	220	13	US-10-424-599-44369	Sequence 44369, A	C 310	14.2	71.0	176	13	US-10-085-783A-17286	Sequence 17286, A
C 238	14.4	72.0	275	17	US-10-437-963-10372	Sequence 10372, A	C 311	14.2	71.0	176	16	US-10-242-535A-17286	Sequence 17286, A
C 239	14.4	72.0	353	13	US-10-424-599-99713	Sequence 99713, A	C 312	14.2	71.0	190	13	US-10-085-783A-22589	Sequence 22589, A
C 240	14.4	72.0	392	10	US-09-918-995-4410	Sequence 4410, Ap	C 313	14.2	71.0	190	16	US-10-242-535A-22589	Sequence 22589, A
C 241	14.4	72.0	487	13	US-10-424-599-10890	Sequence 10890, A	C 314	14.2	71.0	209	9	US-09-960-352-8245	Sequence 8245, Ap
C 242	14.4	72.0	576	9	US-09-777-564-121	Sequence 121, App	C 315	14.2	71.0	218	15	US-10-106-698-1739	Sequence 1739, Ap
C 243	14.4	72.0	576	15	US-10-015-219-121	Sequence 121, App	C 316	14.2	71.0	225	17	US-10-437-963-16934	Sequence 16934, A
C 244	14.4	72.0	582	9	US-09-777-564-555	Sequence 555, App	C 317	14.2	71.0	227	13	US-10-085-783A-9450	Sequence 9450, Ap
C 245	14.4	72.0	582	15	US-10-015-219-555	Sequence 555, App	C 318	14.2	71.0	227	16	US-10-242-535A-9450	Sequence 9450, Ap
C 246	14.4	72.0	594	13	US-10-027-632-96747	Sequence 96747, A	C 319	14.2	71.0	234	15	US-10-029-986A-14659	Sequence 14659, A
C 247	14.4	72.0	594	13	US-10-027-632-306306	Sequence 306306, A	C 320	14.2	71.0	234	16	US-10-393-840-809	Sequence 809, App
C 248	14.4	72.0	594	16	US-10-027-632-96747	Sequence 96747, A	C 321	14.2	71.0	247	9	US-09-864-761-20089	Sequence 20089, A
C 249	14.4	72.0	594	16	US-10-027-632-306306	Sequence 306306, A	C 322	14.2	71.0	249	13	US-10-424-599-96366	Sequence 96366, A
C 250	14.4	72.0	604	17	US-10-021-323-15825	Sequence 15825, A	C 323	14.2	71.0	255	10	US-09-930-213-768	Sequence 768, App
C 251	14.4	72.0	604	13	US-10-027-632-208171	Sequence 208171, A	C 324	14.2	71.0	273	13	US-10-085-783A-13656	Sequence 13656, A
C 252	14.4	72.0	608	13	US-10-027-632-208172	Sequence 208172, A	C 325	14.2	71.0	273	16	US-10-242-535A-13656	Sequence 13656, A
C 253	14.4	72.0	608	13	US-10-027-632-208173	Sequence 208173, A	C 326	14.2	71.0	298	9	US-09-960-352-7655	Sequence 7655, Ap
C 254	14.4	72.0	608	13	US-10-027-632-208174	Sequence 208174, A	C 327	14.2	71.0	298	13	US-10-424-599-140679	Sequence 140679, A
C 255	14.4	72.0	608	13	US-10-027-632-208175	Sequence 208175, A	C 328	14.2	71.0	305	13	US-10-424-599-4051	Sequence 4051, Ap
C 256	14.4	72.0	608	13	US-10-027-632-208176	Sequence 208176, A	C 329	14.2	71.0	315	12	US-09-922-293-1129	Sequence 1129, Ap
C 257	14.4	72.0	608	16	US-10-027-632-208171	Sequence 208171, A	C 330	14.2	71.0	329	8	US-08-781-9866A-986	Sequence 986, App
C 258	14.4	72.0	608	16	US-10-027-632-208172	Sequence 208172, A	C 331	14.2	71.0	329	13	US-10-329-624-886	Sequence 886, App
C 259	14.4	72.0	608	16	US-10-027-632-208173	Sequence 208173, A	C 332	14.2	71.0	330	9	US-09-764-869-445	Sequence 445, App
C 260	14.4	72.0	608	16	US-10-027-632-208174	Sequence 208174, A	C 333	14.2	71.0	330	9	US-09-764-869-445	Sequence 445, App
C 261	14.4	72.0	608	16	US-10-027-632-208175	Sequence 208175, A	C 334	14.2	71.0	330	10	US-09-969-842-71	Sequence 71, Appl
C 262	14.4	72.0	608	16	US-10-027-632-208176	Sequence 208176, A	C 335	14.2	71.0	330	15	US-10-091-504-445	Sequence 445, App
C 263	14.4	72.0	636	13	US-10-027-632-254270	Sequence 254270, A	C 336	14.2	71.0	330	16	US-10-027-577-445	Sequence 445, App
C 264	14.4	72.0	636	16	US-10-027-632-254270	Sequence 254270, A	C 337	14.2	71.0	333	13	US-10-424-599-124672	Sequence 124672, A
C 265	14.4	72.0	654	13	US-10-027-632-217620	Sequence 217620, A	C 338	14.2	71.0	339	15	US-10-029-986-32457	Sequence 32457, A
C 266	14.4	72.0	654	13	US-10-027-632-217621	Sequence 217621, A	C 339	14.2	71.0	341	13	US-10-424-599-11581	Sequence 11581, A
C 267	14.4	72.0	654	16	US-10-027-632-217622	Sequence 217622, A	C 340	14.2	71.0	346	13	US-10-424-599-138636	Sequence 138636, A
C 268	14.4	72.0	654	16	US-10-027-632-217620	Sequence 217620, A	C 341	14.2	71.0	349	13	US-10-424-599-9746	Sequence 9746, Ap
C 269	14.4	72.0	654	16	US-10-027-632-217621	Sequence 217621, A	C 342	14.2	71.0	360	13	US-10-085-783A-48729	Sequence 48729, A
C 270	14.4	72.0	654	16	US-10-027-632-217622	Sequence 217622, A	C 343	14.2	71.0	360	16	US-10-242-535A-48729	Sequence 48729, A
C 271	14.4	72.0	659	17	US-10-021-323-6838	Sequence 6838, Ap	C 344	14.2	71.0	363	13	US-10-424-599-73201	Sequence 73201, A
C 272	14.4	72.0	1155	9	US-09-810-936-329	Sequence 329, App	C 345	14.2	71.0	370	13	US-10-085-783A-110	Sequence 110, App
C 273	14.4	72.0	1155	9	US-09-924-400-329	Sequence 329, App	C 346	14.2	71.0	370	16	US-10-242-535A-110	Sequence 110, App
C 274	14.4	72.0	1155	12	US-10-079-137B-329	Sequence 329, App	C 347	14.2	71.0	373	13	US-10-085-783A-57310	Sequence 57310, A
C 275	14.4	72.0	1155	12	US-10-212-679-329	Sequence 329, App	C 348	14.2	71.0	373	16	US-10-242-535A-57310	Sequence 57310, A
C 276	14.4	72.0	1155	17	US-10-451-168-37	Sequence 37, Appl	C 349	14.2	71.0	376	13	US-10-424-599-5095	Sequence 5095, Ap
C 277	14.4	72.0	1226	13	US-10-027-632-252975	Sequence 252975, A	C 350	14.2	71.0	376	13	US-10-085-783A-25270	Sequence 25270, A
C 278	14.4	72.0	1226	16	US-10-027-632-252975	Sequence 252975, A	C 351	14.2	71.0	376	16	US-10-242-535A-25270	Sequence 25270, A
C 279	14.4	72.0	1233	13	US-10-425-114-20605	Sequence 20605, A	C 352	14.2	71.0	381	13	US-10-282-122A-36660	Sequence 36660, A
C 280	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 353	14.2	71.0	382	9	US-09-960-352-11499	Sequence 11499, A
C 281	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 354	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 282	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 355	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 283	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 356	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 284	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 357	14.2	71.0	384	16	US-10-242-535A-23181	Sequence 23181, A
C 285	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 358	14.2	71.0	400	13	US-10-085-783A-21844	Sequence 21844, A
C 286	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 359	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 287	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 360	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 288	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 361	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 289	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 362	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 290	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 363	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 291	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 364	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 292	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 365	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 293	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 366	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 294	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 367	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 295	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 368	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 296	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 369	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 297	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 370	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 298	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 371	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 299	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 372	14.2	71.0	400	16	US-10-242-535A-21844	Sequence 21844, A
C 300	14.2	71.0	20	16	US-10-199-199-54	Sequence 54, Appl1	C 373	14.2	71.0	442	13	US-10-085-783A-56594	Sequence 56594, A
C 301	14.2	71.0	65	10	US-10-199-199-123	Sequence 123, App	C 374	14.2	71.0	450	17	US-10-242-535A-56594	Sequence 56594, A
C 302	14.2	71.0	130	15	US-09-908-975-24451	Sequence 24451, A	C 375	14.2	71.0	457	17	US-10-437-963-33014	Sequence 33014, A
C 303	14.2	71.0	160	9	US-10-029-386-17205	Sequence 17205, A	C 376	14.2	71.0	460	13	US-10-437-963-78936	Sequence 78936, A
C 304	14.2	71.0	165	9	US-09-864-761-31676	Sequence 31676, A	C 377	14.2	71.0	460	16	US-10-085-783A-25991	Sequence 25991, A
C 305	14.2	71.0	168	15	US-09-864-761-23502	Sequence 23502, A	C 378	14.2	71.0	468	13	US-10-242-535A-25991	Sequence 25991, A
C 306	14.2	71.0	170	9	US-10-029-386-15999	Sequence 15999, A	C 379	14.2	71.0	468	16	US-10-085-783A-26172	Sequence 26172, A
					Sequence 151, App								

C 380	14.2	71.0	470	13	US-10-085-783A-42081	Sequence 42081, A	C 453	14.2	71.0	828	16	US-10-027-632-173547	Sequence 173547, A
C 381	14.2	71.0	470	16	US-10-248-535A-42081	Sequence 42081, A	C 454	14.2	71.0	828	16	US-10-027-632-173548	Sequence 173548, A
C 382	14.2	71.0	471	9	US-09-783-590-10451	Sequence 10451, A	C 455	14.2	71.0	828	16	US-10-027-632-173549	Sequence 173549, A
C 383	14.2	71.0	473	9	US-09-864-761-15149	Sequence 15149, A	C 456	14.2	71.0	842	13	US-10-425-114-21555	Sequence 21555, A
C 384	14.2	71.0	476	13	US-10-027-632-181465	Sequence 181465, A	C 457	14.2	71.0	870	13	US-10-282-122A-27655	Sequence 27655, A
C 385	14.2	71.0	476	13	US-10-027-632-181466	Sequence 181466, A	C 458	14.2	71.0	884	16	US-10-369-493-30485	Sequence 30485, A
C 386	14.2	71.0	476	16	US-10-027-632-181465	Sequence 181465, A	C 459	14.2	71.0	912	13	US-10-424-599-26124	Sequence 26124, A
C 387	14.2	71.0	476	16	US-10-027-632-181466	Sequence 181466, A	C 460	14.2	71.0	915	13	US-10-282-122A-39868	Sequence 39868, A
C 388	14.2	71.0	477	10	US-09-918-995-12069	Sequence 12069, A	C 461	14.2	71.0	928	13	US-10-425-114-19003	Sequence 19003, A
C 389	14.2	71.0	482	10	US-09-918-995-33549	Sequence 33549, A	C 462	14.2	71.0	928	13	US-10-282-122A-27731	Sequence 27731, A
C 390	14.2	71.0	484	9	US-09-864-761-3314	Sequence 3314, A	C 463	14.2	71.0	934	13	US-10-425-114-20130	Sequence 20130, A
C 391	14.2	71.0	486	9	US-09-864-761-5456	Sequence 5456, A	C 464	14.2	71.0	936	13	US-10-425-114-3330	Sequence 3330, A
C 392	14.2	71.0	506	9	US-09-796-692-2838	Sequence 2838, A	C 465	14.2	71.0	936	13	US-10-425-114-14187	Sequence 14187, A
C 393	14.2	71.0	506	15	US-10-040-862-2838	Sequence 2838, A	C 466	14.2	71.0	954	13	US-10-425-114-29727	Sequence 29727, A
C 394	14.2	71.0	506	16	US-10-057-475B-2838	Sequence 2838, A	C 467	14.2	71.0	970	13	US-10-342-887-2667	Sequence 2667, A
C 395	14.2	71.0	506	16	US-10-155-884B-2838	Sequence 2838, A	C 468	14.2	71.0	970	13	US-10-172-118-2667	Sequence 2667, A
C 396	14.2	71.0	507	15	US-10-028-386-9845	Sequence 9845, A	C 469	14.2	71.0	993	15	US-10-156-761-4170	Sequence 4170, A
C 397	14.2	71.0	521	13	US-10-027-632-312912	Sequence 312912, A	C 470	14.2	71.0	1008	17	US-10-437-563-26282	Sequence 26282, A
C 398	14.2	71.0	521	16	US-10-027-632-312912	Sequence 312912, A	C 471	14.2	71.0	1018	13	US-10-027-632-10345	Sequence 10345, A
C 399	14.2	71.0	534	15	US-10-028-386-13636	Sequence 13636, A	C 472	14.2	71.0	1018	13	US-10-027-632-10345	Sequence 10345, A
C 400	14.2	71.0	548	9	US-09-864-761-6769	Sequence 6769, A	C 473	14.2	71.0	1018	16	US-10-027-632-10346	Sequence 10346, A
C 401	14.2	71.0	548	13	US-10-425-114-16250	Sequence 16250, A	C 474	14.2	71.0	1018	16	US-10-027-632-10346	Sequence 10346, A
C 402	14.2	71.0	548	15	US-10-028-386-850	Sequence 850, A	C 475	14.2	71.0	1025	9	US-09-971-536-25	Sequence 25, A
C 403	14.2	71.0	549	15	US-10-028-386-956	Sequence 956, A	C 476	14.2	71.0	1045	9	US-10-425-114-17568	Sequence 17568, A
C 404	14.2	71.0	554	17	US-10-437-963-73593	Sequence 73593, A	C 477	14.2	71.0	1059	13	US-10-425-114-11667	Sequence 11667, A
C 405	14.2	71.0	558	15	US-10-028-386-2299	Sequence 2299, A	C 478	14.2	71.0	1098	13	US-10-425-114-5886	Sequence 5886, A
C 406	14.2	71.0	560	15	US-10-028-386-3505	Sequence 3505, A	C 479	14.2	71.0	1104	13	US-10-424-599-35651	Sequence 35651, A
C 407	14.2	71.0	567	13	US-10-240-425-439	Sequence 439, A	C 480	14.2	71.0	1106	13	US-10-027-632-117171	Sequence 117171, A
C 408	14.2	71.0	567	13	US-10-027-632-49222	Sequence 49222, A	C 481	14.2	71.0	1106	13	US-10-027-632-117171	Sequence 117171, A
C 409	14.2	71.0	568	13	US-10-027-632-63505	Sequence 63505, A	C 482	14.2	71.0	1106	16	US-10-027-632-117171	Sequence 117171, A
C 410	14.2	71.0	568	13	US-10-027-632-88517	Sequence 88517, A	C 483	14.2	71.0	1106	16	US-10-027-632-117172	Sequence 117172, A
C 411	14.2	71.0	568	16	US-10-027-632-49222	Sequence 49222, A	C 484	14.2	71.0	1116	13	US-10-282-122A-20688	Sequence 20688, A
C 412	14.2	71.0	568	16	US-10-027-632-63505	Sequence 63505, A	C 485	14.2	71.0	1148	13	US-10-225-066A-991	Sequence 991, A
C 413	14.2	71.0	568	16	US-10-027-632-88517	Sequence 88517, A	C 486	14.2	71.0	1148	16	US-10-374-780A-2813	Sequence 2813, A
C 414	14.2	71.0	589	13	US-10-027-632-24849	Sequence 24849, A	C 487	14.2	71.0	1156	13	US-10-027-632-210835	Sequence 210835, A
C 415	14.2	71.0	589	13	US-10-027-632-248500	Sequence 248500, A	C 488	14.2	71.0	1156	16	US-10-027-632-210835	Sequence 210835, A
C 416	14.2	71.0	589	13	US-10-027-632-248501	Sequence 248501, A	C 489	14.2	71.0	1176	10	US-09-729-658B-14	Sequence 14, A
C 417	14.2	71.0	589	16	US-10-027-632-248499	Sequence 248499, A	C 490	14.2	71.0	1187	13	US-10-424-599-320843	Sequence 320843, A
C 418	14.2	71.0	589	16	US-10-027-632-248500	Sequence 248500, A	C 491	14.2	71.0	1206	16	US-10-369-493-32245	Sequence 32245, A
C 419	14.2	71.0	589	16	US-10-027-632-248501	Sequence 248501, A	C 492	14.2	71.0	1214	9	US-09-764-846-52	Sequence 52, A
C 420	14.2	71.0	615	17	US-10-437-963-63535	Sequence 63535, A	C 493	14.2	71.0	1214	15	US-10-091-483-52	Sequence 52, A
C 421	14.2	71.0	622	13	US-10-027-632-17865	Sequence 17865, A	C 494	14.2	71.0	1248	13	US-10-282-122A-36327	Sequence 36327, A
C 422	14.2	71.0	622	13	US-10-027-632-178654	Sequence 178654, A	C 495	14.2	71.0	1254	17	US-10-437-963-43994	Sequence 43994, A
C 423	14.2	71.0	622	16	US-10-027-632-178685	Sequence 178685, A	C 496	14.2	71.0	1271	13	US-10-027-632-121256	Sequence 121256, A
C 424	14.2	71.0	622	16	US-10-027-632-178694	Sequence 178694, A	C 497	14.2	71.0	1271	16	US-10-027-632-121256	Sequence 121256, A
C 425	14.2	71.0	627	13	US-10-027-632-5709	Sequence 5709, A	C 498	14.2	71.0	1281	13	US-10-282-122A-27814	Sequence 27814, A
C 426	14.2	71.0	627	16	US-10-027-632-5709	Sequence 5709, A	C 499	14.2	71.0	1287	13	US-10-282-122A-18884	Sequence 18884, A
C 427	14.2	71.0	637	13	US-10-027-632-190606	Sequence 190606, A	C 500	14.2	71.0	1290	9	US-09-764-846-124	Sequence 124, A
C 428	14.2	71.0	637	13	US-10-027-632-230816	Sequence 230816, A	C 501	14.2	71.0	1290	15	US-10-091-483-124	Sequence 124, A
C 429	14.2	71.0	637	16	US-10-027-632-190606	Sequence 190606, A	C 502	14.2	71.0	1314	16	US-10-369-493-24743	Sequence 24743, A
C 430	14.2	71.0	637	16	US-10-027-632-230816	Sequence 230816, A	C 503	14.2	71.0	1368	13	US-10-424-599-121015	Sequence 121015, A
C 431	14.2	71.0	643	13	US-10-027-632-238961	Sequence 238961, A	C 504	14.2	71.0	1407	9	US-10-425-114-20301	Sequence 20301, A
C 432	14.2	71.0	643	13	US-10-027-632-238962	Sequence 238962, A	C 505	14.2	71.0	1452	9	US-09-938-842A-2402	Sequence 2402, A
C 433	14.2	71.0	643	13	US-10-027-632-238963	Sequence 238963, A	C 506	14.2	71.0	1452	11	US-09-938-842A-2402	Sequence 2402, A
C 434	14.2	71.0	643	16	US-10-027-632-238961	Sequence 238961, A	C 507	14.2	71.0	1467	17	US-10-437-963-33825	Sequence 33825, A
C 435	14.2	71.0	643	16	US-10-027-632-238962	Sequence 238962, A	C 508	14.2	71.0	1470	9	US-09-938-842A-1796	Sequence 1796, A
C 436	14.2	71.0	644	16	US-10-027-632-238963	Sequence 238963, A	C 509	14.2	71.0	1470	11	US-09-938-842A-1796	Sequence 1796, A
C 437	14.2	71.0	644	13	US-10-027-632-94592	Sequence 94592, A	C 510	14.2	71.0	1574	10	US-09-729-658B-1	Sequence 1, A
C 438	14.2	71.0	644	13	US-10-027-632-305513	Sequence 305513, A	C 511	14.2	71.0	1591	16	US-10-320-197-1159	Sequence 1159, A
C 439	14.2	71.0	644	16	US-10-027-632-94592	Sequence 94592, A	C 512	14.2	71.0	1604	13	US-10-425-114-28650	Sequence 28650, A
C 440	14.2	71.0	644	16	US-10-027-632-305513	Sequence 305513, A	C 513	14.2	71.0	1637	13	US-10-425-114-20378	Sequence 20378, A
C 441	14.2	71.0	685	13	US-10-027-632-113381	Sequence 113381, A	C 514	14.2	71.0	1647	13	US-10-424-599-50848	Sequence 50848, A
C 442	14.2	71.0	685	16	US-10-027-632-113381	Sequence 113381, A	C 515	14.2	71.0	1651	10	US-09-533-029-59	Sequence 59, A
C 443	14.2	71.0	703	17	US-10-437-963-17503	Sequence 17503, A	C 516	14.2	71.0	1651	13	US-10-412-699B-881	Sequence 881, A
C 444	14.2	71.0	725	16	US-10-027-632-20217	Sequence 20217, A	C 517	14.2	71.0	1651	13	US-10-425-114-20130	Sequence 20130, A
C 445	14.2	71.0	743	10	US-09-729-658B-11	Sequence 11, A	C 518	14.2	71.0	1651	16	US-10-374-780A-271	Sequence 271, A
C 446	14.2	71.0	752	9	US-09-822-849A-556	Sequence 556, A	C 519	14.2	71.0	1653	13	US-10-282-122A-27624	Sequence 27624, A
C 447	14.2	71.0	752	13	US-10-027-632-150015	Sequence 150015, A	C 520	14.2	71.0	1654	16	US-10-264-413-16	Sequence 16, A
C 448	14.2	71.0	752	16	US-10-027-632-150015	Sequence 150015, A	C 521	14.2	71.0	1725	13	US-10-282-122A-30934	Sequence 30934, A
C 449	14.2	71.0	826	13	US-10-424-599-135108	Sequence 135108, A	C 522	14.2	71.0	1745	13	US-10-424-599-38057	Sequence 38057, A
C 450	14.2	71.0	828	13	US-10-027-632-173547	Sequence 173547, A	C 523	14.2	71.0	1745	13	US-10-425-114-5089	Sequence 5089, A
C 451	14.2	71.0	828	13	US-10-027-632-173548	Sequence 173548, A	C 524	14.2	71.0	1747	13	US-10-425-114-19556	Sequence 19556, A
C 452	14.2	71.0	828	13	US-10-027-632-173549	Sequence 173549, A	C 525	14.2	71.0	1790	10	US-09-919-039-150	Sequence 150, A

C 526	14.2	71.0	1790	14	US-10-044-090-505	Sequence 505, App	C 599	14.2	71.0	3169	9	US-09-981-353-53	Sequence 53, App1
C 527	14.2	71.0	1838	13	US-10-425-114-5743	Sequence 573, Ap	C 600	14.2	71.0	3169	15	US-10-235-994-15	Sequence 15, App1
C 528	14.2	71.0	1842	16	US-10-429-873A-1	Sequence 1, App1	C 601	14.2	71.0	3181	14	US-10-025-167-18	Sequence 18, App1
C 529	14.2	71.0	1861	13	US-10-425-114-620	Sequence 620, App	C 602	14.2	71.0	3188	16	US-10-108-260A-581	Sequence 581, App
C 530	14.2	71.0	1879	17	US-10-647-426-30	Sequence 30, App1	C 603	14.2	71.0	3195	10	US-09-867-034-22	Sequence 22, App1
C 531	14.2	71.0	1883	15	US-10-210-120-98	Sequence 98, App1	C 604	14.2	71.0	3195	13	US-10-276-115-22	Sequence 22, App1
C 532	14.2	71.0	1890	16	US-10-159-563-390	Sequence 390, App	C 605	14.2	71.0	3196	15	US-10-158-646-39	Sequence 39, App1
C 533	14.2	71.0	1899	13	US-10-282-122A-9684	Sequence 9684, Ap	C 606	14.2	71.0	3199	13	US-10-276-774-93	Sequence 93, App1
C 534	14.2	71.0	1902	16	US-10-398-221-1054	Sequence 1054, Ap	C 607	14.2	71.0	3204	15	US-10-345-680-31	Sequence 31, App1
C 535	14.2	71.0	1902	16	US-10-398-221-2831	Sequence 2831, Ap	C 608	14.2	71.0	3207	15	US-10-101-510-60	Sequence 60, App
C 536	14.2	71.0	1905	13	US-10-425-114-3032	Sequence 32032, A	C 609	14.2	71.0	3218	16	US-10-087-080-33	Sequence 33, App1
C 537	14.2	71.0	1908	13	US-10-282-122A-14840	Sequence 14840, A	C 610	14.2	71.0	3234	13	US-10-423-543-8	Sequence 48, App1
C 538	14.2	71.0	1954	17	US-10-641-643-286	Sequence 286, App	C 611	14.2	71.0	3262	13	US-10-369-922-57	Sequence 57, App1
C 539	14.2	71.0	1962	15	US-10-233-665-5	Sequence 5, App1	C 612	14.2	71.0	3262	16	US-10-199-199-4	Sequence 4, App1
C 540	14.2	71.0	1989	10	US-09-738-630-93	Sequence 93, App1	C 613	14.2	71.0	3265	9	US-09-989-722-378	Sequence 378, App
C 541	14.2	71.0	1989	15	US-10-175-523-190	Sequence 190, App	C 614	14.2	71.0	3265	9	US-09-989-723-378	Sequence 378, App
C 542	14.2	71.0	1989	15	US-10-102-524-1725	Sequence 1725, Ap	C 615	14.2	71.0	3265	9	US-09-989-729-378	Sequence 378, App
C 543	14.2	71.0	1989	15	US-10-232-665-7	Sequence 7, App1	C 616	14.2	71.0	3265	9	US-09-989-727-378	Sequence 378, App
C 544	14.2	71.0	2000	13	US-10-614-076-99	Sequence 99, App1	C 617	14.2	71.0	3265	9	US-09-989-731-378	Sequence 378, App
C 545	14.2	71.0	2000	16	US-10-260-238-1960	Sequence 1960, Ap	C 618	14.2	71.0	3265	9	US-09-989-732-378	Sequence 378, App
C 546	14.2	71.0	2002	9	US-09-935-390A-5	Sequence 5, App1	C 619	14.2	71.0	3265	9	US-09-991-073-378	Sequence 378, App
C 547	14.2	71.0	2016	16	US-10-369-493-37033	Sequence 37033, A	C 620	14.2	71.0	3265	9	US-09-990-442-378	Sequence 378, App
C 548	14.2	71.0	2050	13	US-10-614-076-101	Sequence 101, App	C 621	14.2	71.0	3265	9	US-09-991-163-378	Sequence 378, App
C 549	14.2	71.0	2082	13	US-10-027-632-98674	Sequence 98674, A	C 622	14.2	71.0	3265	9	US-09-991-604-378	Sequence 378, App
C 550	14.2	71.0	2082	16	US-10-027-632-98674	Sequence 98674, A	C 623	14.2	71.0	3265	9	US-09-990-456-378	Sequence 378, App
C 551	14.2	71.0	2109	9	US-09-925-297-83	Sequence 83, App1	C 624	14.2	71.0	3265	9	US-09-989-721-378	Sequence 378, App
C 552	14.2	71.0	2110	16	US-10-104-047-743	Sequence 743, App	C 625	14.2	71.0	3265	9	US-09-992-558-378	Sequence 378, App
C 553	14.2	71.0	2194	17	US-10-437-963-62181	Sequence 62181, A	C 626	14.2	71.0	3265	9	US-09-989-253A-378	Sequence 378, App
C 554	14.2	71.0	2200	17	US-09-925-301-217	Sequence 217, App	C 627	14.2	71.0	3265	9	US-09-989-735-378	Sequence 378, App
C 555	14.2	71.0	2200	17	US-10-322-281-812	Sequence 812, App	C 628	14.2	71.0	3265	9	US-09-990-444-378	Sequence 378, App
C 556	14.2	71.0	2211	15	US-10-171-581-119	Sequence 119, App	C 629	14.2	71.0	3265	9	US-09-991-181-378	Sequence 378, App
C 557	14.2	71.0	2215	13	US-10-283-122A-33208	Sequence 33208, A	C 630	14.2	71.0	3265	9	US-09-991-178-378	Sequence 378, App
C 558	14.2	71.0	2265	17	US-10-437-963-49826	Sequence 49826, A	C 631	14.2	71.0	3265	9	US-09-990-426-378	Sequence 378, App
C 559	14.2	71.0	2274	13	US-10-282-122A-31568	Sequence 31568, A	C 632	14.2	71.0	3265	9	US-09-993-687-378	Sequence 378, App
C 560	14.2	71.0	2313	16	US-10-369-493-41968	Sequence 41968, A	C 633	14.2	71.0	3265	10	US-09-989-734-378	Sequence 378, App
C 561	14.2	71.0	2344	15	US-10-352-762-35	Sequence 35, App1	C 634	14.2	71.0	3265	10	US-09-993-653-378	Sequence 378, App
C 562	14.2	71.0	2347	15	US-10-352-762-34	Sequence 1, App1	C 635	14.2	71.0	3265	10	US-09-993-657-378	Sequence 378, App
C 563	14.2	71.0	2347	15	US-10-352-762-34	Sequence 34, App1	C 636	14.2	71.0	3265	10	US-09-997-428-378	Sequence 378, App
C 564	14.2	71.0	2347	15	US-10-352-762-36	Sequence 36, App1	C 637	14.2	71.0	3265	10	US-09-997-666-378	Sequence 378, App
C 565	14.2	71.0	2373	9	US-09-964-824A-253	Sequence 253, App	C 638	14.2	71.0	3265	10	US-09-990-438-378	Sequence 378, App
C 566	14.2	71.0	2373	9	US-09-880-107-2365	Sequence 2365, Ap	C 639	14.2	71.0	3265	10	US-09-990-562-378	Sequence 378, App
C 567	14.2	71.0	2373	15	US-10-101-510-745	Sequence 745, App	C 640	14.2	71.0	3265	10	US-09-990-711-378	Sequence 378, App
C 568	14.2	71.0	2373	13	US-10-641-643-1011	Sequence 1011, Ap	C 641	14.2	71.0	3265	10	US-09-989-126-378	Sequence 378, App
C 569	14.2	71.0	2429	14	US-10-116-802-98	Sequence 98, App1	C 642	14.2	71.0	3265	10	US-09-990-437-378	Sequence 378, App
C 570	14.2	71.0	2429	14	US-10-044-090-344	Sequence 344, App	C 643	14.2	71.0	3265	10	US-09-991-157-378	Sequence 378, App
C 571	14.2	71.0	2436	17	US-10-437-963-283	Sequence 283, App	C 644	14.2	71.0	3265	10	US-09-997-514-378	Sequence 378, App
C 572	14.2	71.0	2485	15	US-10-084-817-228	Sequence 228, App	C 645	14.2	71.0	3265	10	US-09-997-514-378	Sequence 378, App
C 573	14.2	71.0	2485	15	US-10-101-510-508	Sequence 508, App	C 646	14.2	71.0	3265	10	US-09-997-514-378	Sequence 378, App
C 574	14.2	71.0	2485	15	US-10-252-157-206	Sequence 206, App	C 647	14.2	71.0	3265	10	US-09-991-172-378	Sequence 378, App
C 575	14.2	71.0	2487	13	US-10-283-132A-32996	Sequence 32996, A	C 648	14.2	71.0	3265	10	US-09-990-726-378	Sequence 378, App
C 576	14.2	71.0	2502	17	US-10-437-963-95565	Sequence 95565, A	C 649	14.2	71.0	3265	10	US-09-997-559-378	Sequence 378, App
C 577	14.2	71.0	2506	12	US-10-041-018-377	Sequence 377, App	C 650	14.2	71.0	3265	10	US-09-997-601-378	Sequence 378, App
C 578	14.2	71.0	2562	16	US-10-369-493-28567	Sequence 28567, A	C 651	14.2	71.0	3265	10	US-09-990-443-378	Sequence 378, App
C 579	14.2	71.0	2574	17	US-10-437-963-34290	Sequence 34290, A	C 652	14.2	71.0	3265	10	US-09-991-854-378	Sequence 378, App
C 580	14.2	71.0	2607	16	US-10-369-493-31325	Sequence 31325, A	C 653	14.2	71.0	3265	10	US-09-997-628-378	Sequence 378, App
C 581	14.2	71.0	2684	17	US-10-283-975A-48	Sequence 48, App1	C 654	14.2	71.0	3265	10	US-09-987-683-378	Sequence 378, App
C 582	14.2	71.0	2754	15	US-10-345-680-33	Sequence 33, App1	C 655	14.2	71.0	3265	10	US-09-989-729A-378	Sequence 378, App
C 583	14.2	71.0	2771	15	US-10-223-567A-468	Sequence 468, App	C 656	14.2	71.0	3265	10	US-09-997-449-378	Sequence 378, App
C 584	14.2	71.0	2825	13	US-10-425-114-42222	Sequence 22222, A	C 657	14.2	71.0	3265	10	US-09-997-440-378	Sequence 378, App
C 585	14.2	71.0	2834	9	US-09-954-456-286	Sequence 286, App	C 658	14.2	71.0	3265	10	US-09-990-440-378	Sequence 378, App
C 586	14.2	71.0	2834	9	US-09-954-456-286	Sequence 954, App	C 659	14.2	71.0	3265	10	US-09-993-469-378	Sequence 378, App
C 587	14.2	71.0	2834	9	US-09-954-456-1603	Sequence 1603, App	C 660	14.2	71.0	3265	10	US-09-997-542-378	Sequence 378, App
C 588	14.2	71.0	2834	16	US-09-873-367C-311	Sequence 311, App	C 661	14.2	71.0	3265	10	US-09-993-748-378	Sequence 378, App
C 589	14.2	71.0	2834	16	US-10-295-027-874	Sequence 874, App	C 662	14.2	71.0	3265	10	US-09-990-439-378	Sequence 378, App
C 590	14.2	71.0	2841	15	US-10-128-714-57	Sequence 57, App1	C 663	14.2	71.0	3265	10	US-09-990-427-378	Sequence 378, App
C 591	14.2	71.0	2841	15	US-10-128-714-57	Sequence 5057, App	C 664	14.2	71.0	3265	10	US-09-989-128-378	Sequence 378, App
C 592	14.2	71.0	2852	16	US-10-094-749-1113	Sequence 1113, Ap	C 665	14.2	71.0	3265	10	US-09-993-583-378	Sequence 378, App
C 593	14.2	71.0	2867	13	US-10-116-802-155	Sequence 155, App	C 666	14.2	71.0	3265	10	US-09-941-892-378	Sequence 378, App
C 594	14.2	71.0	2981	15	US-10-096-534-45	Sequence 45, App1	C 667	14.2	71.0	3265	10	US-09-992-521-378	Sequence 378, App
C 595	14.2	71.0	3003	16	US-10-429-873A-5	Sequence 5, App1	C 668	14.2	71.0	3265	10	US-09-997-333-378	Sequence 378, App
C 596	14.2	71.0	3003	16	US-10-429-873A-6	Sequence 6, App1	C 669	14.2	71.0	3265	10	US-09-997-384-378	Sequence 378, App
C 597	14.2	71.0	3043	14	US-10-025-167-16	Sequence 16, App1	C 670	14.2	71.0	3265	10	US-09-998-041-378	Sequence 378, App
C 598	14.2	71.0	3044	15	US-10-233-665-38	Sequence 38, App1	C 671	14.2	71.0	3265	10	US-09-997-585-378	Sequence 378, App

C 672	14.2	71.0	3265	10	US-09-997-614-378	Sequence 378, App
C 673	14.2	71.0	3265	10	US-09-989-862-378	Sequence 378, App
C 674	14.2	71.0	3265	10	US-09-997-529-378	Sequence 378, App
C 675	14.2	71.0	3265	10	US-09-989-725-378	Sequence 378, App
C 676	14.2	71.0	3265	11	US-09-989-733-378	Sequence 378, App
C 677	14.2	71.0	3265	11	US-09-992-643-378	Sequence 378, App
C 678	14.2	71.0	3265	13	US-10-206-915-257	Sequence 257, App
C 679	14.2	71.0	3265	13	US-10-199-670-257	Sequence 257, App
C 680	14.2	71.0	3265	13	US-10-201-858-257	Sequence 257, App
C 681	14.2	71.0	3265	13	US-10-205-890-257	Sequence 257, App
C 682	14.2	71.0	3265	13	US-10-208-024-257	Sequence 257, App
C 683	14.2	71.0	3265	13	US-10-201-853-257	Sequence 257, App
C 684	14.2	71.0	3265	13	US-10-063-745-69	Sequence 69, Appl
C 685	14.2	71.0	3265	13	US-09-989-724-378	Sequence 378, App
C 686	14.2	71.0	3265	13	US-09-989-728-378	Sequence 378, App
C 687	14.2	71.0	3265	13	US-09-990-441-378	Sequence 378, App
C 688	14.2	71.0	3265	13	US-10-063-512-69	Sequence 69, Appl
C 689	14.2	71.0	3265	13	US-10-063-513-69	Sequence 69, Appl
C 690	14.2	71.0	3265	13	US-10-063-515-69	Sequence 69, Appl
C 691	14.2	71.0	3265	13	US-10-063-549-69	Sequence 69, Appl
C 692	14.2	71.0	3265	13	US-10-063-569-69	Sequence 69, Appl
C 693	14.2	71.0	3265	13	US-10-063-581-69	Sequence 69, Appl
C 694	14.2	71.0	3265	13	US-10-174-581-257	Sequence 257, App
C 695	14.2	71.0	3265	13	US-10-176-483-257	Sequence 257, App
C 696	14.2	71.0	3265	13	US-10-176-749-257	Sequence 257, App
C 697	14.2	71.0	3265	13	US-10-176-914-257	Sequence 257, App
C 698	14.2	71.0	3265	13	US-10-176-915-257	Sequence 257, App
C 699	14.2	71.0	3265	13	US-09-997-857-378	Sequence 378, App
C 700	14.2	71.0	3265	13	US-10-063-555-69	Sequence 69, Appl
C 701	14.2	71.0	3265	13	US-10-063-563-69	Sequence 69, Appl
C 702	14.2	71.0	3265	13	US-10-063-594-69	Sequence 69, Appl
C 703	14.2	71.0	3265	13	US-10-063-553-69	Sequence 69, Appl
C 704	14.2	71.0	3265	13	US-10-063-554-69	Sequence 69, Appl
C 705	14.2	71.0	3265	13	US-10-176-484-257	Sequence 257, App
C 706	14.2	71.0	3265	13	US-10-180-550-257	Sequence 257, App
C 707	14.2	71.0	3265	13	US-10-183-014-257	Sequence 257, App
C 708	14.2	71.0	3265	13	US-10-187-738-257	Sequence 257, App
C 709	14.2	71.0	3265	13	US-10-187-740-257	Sequence 257, App
C 710	14.2	71.0	3265	13	US-10-187-883-257	Sequence 257, App
C 711	14.2	71.0	3265	13	US-10-194-363-257	Sequence 257, App
C 712	14.2	71.0	3265	13	US-10-194-460-257	Sequence 257, App
C 713	14.2	71.0	3265	13	US-10-194-463-257	Sequence 257, App
C 714	14.2	71.0	3265	13	US-10-194-484-257	Sequence 257, App
C 715	14.2	71.0	3265	13	US-10-195-884-257	Sequence 257, App
C 716	14.2	71.0	3265	13	US-10-195-896-257	Sequence 257, App
C 717	14.2	71.0	3265	13	US-10-196-744-257	Sequence 257, App
C 718	14.2	71.0	3265	13	US-10-196-755-257	Sequence 257, App
C 719	14.2	71.0	3265	13	US-10-196-757-257	Sequence 257, App
C 720	14.2	71.0	3265	13	US-10-197-704-257	Sequence 257, App
C 721	14.2	71.0	3265	13	US-10-197-710-257	Sequence 257, App
C 722	14.2	71.0	3265	13	US-10-198-758-257	Sequence 257, App
C 723	14.2	71.0	3265	13	US-10-198-766-257	Sequence 257, App
C 724	14.2	71.0	3265	13	US-10-199-304-257	Sequence 257, App
C 725	14.2	71.0	3265	13	US-10-199-309-257	Sequence 257, App
C 726	14.2	71.0	3265	13	US-10-199-313-257	Sequence 257, App
C 727	14.2	71.0	3265	13	US-10-199-456-257	Sequence 257, App
C 728	14.2	71.0	3265	13	US-10-201-329-257	Sequence 257, App
C 729	14.2	71.0	3265	13	US-10-203-412-257	Sequence 257, App
C 730	14.2	71.0	3265	13	US-10-206-919-257	Sequence 257, App
C 731	14.2	71.0	3265	13	US-10-206-922-257	Sequence 257, App
C 732	14.2	71.0	3265	13	US-10-206-924-257	Sequence 257, App
C 733	14.2	71.0	3265	13	US-10-206-928-257	Sequence 257, App
C 734	14.2	71.0	3265	13	US-10-207-914-257	Sequence 257, App
C 735	14.2	71.0	3265	13	US-10-207-921-257	Sequence 257, App
C 736	14.2	71.0	3265	13	US-10-207-922-257	Sequence 257, App
C 737	14.2	71.0	3265	13	US-10-208-027-257	Sequence 257, App
C 738	14.2	71.0	3265	13	US-09-997-641-378	Sequence 378, App
C 739	14.2	71.0	3265	13	US-09-991-150-378	Sequence 378, App
C 740	14.2	71.0	3265	13	US-10-174-570-257	Sequence 257, App
C 741	14.2	71.0	3265	13	US-10-183-005-257	Sequence 257, App
C 742	14.2	71.0	3265	14	US-10-006-867-69	Sequence 69, Appl
C 743	14.2	71.0	3265	14	US-10-052-586-257	Sequence 257, App
C 744	14.2	71.0	3265	14	US-10-063-547-69	Sequence 69, Appl
C 745	14.2	71.0	3265	15	US-10-174-590-257	Sequence 257, App
C 746	14.2	71.0	3265	15	US-10-176-758-257	Sequence 257, App
C 747	14.2	71.0	3265	15	US-10-175-737-257	Sequence 257, App
C 748	14.2	71.0	3265	15	US-10-063-616-69	Sequence 69, Appl
C 749	14.2	71.0	3265	15	US-10-173-706-69	Sequence 257, App
C 750	14.2	71.0	3265	15	US-10-173-738-257	Sequence 257, App
C 751	14.2	71.0	3265	15	US-10-175-752-257	Sequence 257, App
C 752	14.2	71.0	3265	15	US-10-176-482-257	Sequence 257, App
C 753	14.2	71.0	3265	15	US-10-176-757-257	Sequence 257, App
C 754	14.2	71.0	3265	15	US-10-176-813-257	Sequence 257, App
C 755	14.2	71.0	3265	15	US-10-180-552-257	Sequence 257, App
C 756	14.2	71.0	3265	15	US-10-180-557-257	Sequence 257, App
C 757	14.2	71.0	3265	15	US-10-063-502-69	Sequence 69, Appl
C 758	14.2	71.0	3265	15	US-10-173-700-257	Sequence 257, App
C 759	14.2	71.0	3265	15	US-10-174-572-257	Sequence 257, App
C 760	14.2	71.0	3265	15	US-10-174-579-257	Sequence 257, App
C 761	14.2	71.0	3265	15	US-10-174-582-257	Sequence 257, App
C 762	14.2	71.0	3265	15	US-10-174-588-257	Sequence 257, App
C 763	14.2	71.0	3265	15	US-10-175-739-257	Sequence 257, App
C 764	14.2	71.0	3265	15	US-10-175-740-257	Sequence 257, App
C 765	14.2	71.0	3265	15	US-10-175-743-257	Sequence 257, App
C 766	14.2	71.0	3265	15	US-10-176-488-257	Sequence 257, App
C 767	14.2	71.0	3265	15	US-10-176-492-257	Sequence 257, App
C 768	14.2	71.0	3265	15	US-10-176-747-257	Sequence 257, App
C 769	14.2	71.0	3265	15	US-10-176-750-257	Sequence 257, App
C 770	14.2	71.0	3265	15	US-10-176-985-257	Sequence 257, App
C 771	14.2	71.0	3265	15	US-10-176-987-257	Sequence 257, App
C 772	14.2	71.0	3265	15	US-10-176-992-257	Sequence 257, App
C 773	14.2	71.0	3265	15	US-10-176-993-257	Sequence 257, App
C 774	14.2	71.0	3265	15	US-10-184-658-257	Sequence 257, App
C 775	14.2	71.0	3265	15	US-10-176-991-257	Sequence 257, App
C 776	14.2	71.0	3265	15	US-10-173-695-257	Sequence 257, App
C 777	14.2	71.0	3265	15	US-10-173-697-257	Sequence 257, App
C 778	14.2	71.0	3265	15	US-10-173-705-257	Sequence 257, App
C 779	14.2	71.0	3265	15	US-10-174-576-257	Sequence 257, App
C 780	14.2	71.0	3265	15	US-10-174-585-257	Sequence 257, App
C 781	14.2	71.0	3265	15	US-10-174-586-257	Sequence 257, App
C 782	14.2	71.0	3265	15	US-10-175-747-257	Sequence 257, App
C 783	14.2	71.0	3265	15	US-10-176-481-257	Sequence 257, App
C 784	14.2	71.0	3265	15	US-10-176-485-257	Sequence 257, App
C 785	14.2	71.0	3265	15	US-10-176-487-257	Sequence 257, App
C 786	14.2	71.0	3265	15	US-10-176-493-257	Sequence 257, App
C 787	14.2	71.0	3265	15	US-10-176-756-257	Sequence 257, App
C 788	14.2	71.0	3265	15	US-10-176-911-257	Sequence 257, App
C 789	14.2	71.0	3265	15	US-10-176-919-257	Sequence 257, App
C 790	14.2	71.0	3265	15	US-10-176-925-257	Sequence 257, App
C 791	14.2	71.0	3265	15	US-10-176-928-257	Sequence 257, App
C 792	14.2	71.0	3265	15	US-10-179-510-257	Sequence 257, App
C 793	14.2	71.0	3265	15	US-10-180-543-257	Sequence 257, App
C 794	14.2	71.0	3265	15	US-10-180-544-257	Sequence 257, App
C 795	14.2	71.0	3265	15	US-10-180-546-257	Sequence 257, App
C 796	14.2	71.0	3265	15	US-10-180-547-257	Sequence 257, App
C 797	14.2	71.0	3265	15	US-10-180-549-257	Sequence 257, App
C 798	14.2	71.0	3265	15	US-10-180-555-257	Sequence 257, App
C 799	14.2	71.0	3265	15	US-10-180-559-257	Sequence 257, App
C 800	14.2	71.0	3265	15	US-10-181-000-257	Sequence 257, App

ALIGNMENTS

RESULT 1

US-09-825-489-3

Sequence 3, Application US/09825489

Publication No. US20030232767A1

GENERAL INFORMATION:

APPLICANT: AGRAWAL, SUDHIR

APPLICANT: KANDIMALLA, EKAMBAR R.

APPLICANT: BREGMAN, DAVID B.

APPLICANT: MANT, SRIDHAR

APPLICANT: LU, YI

TITLE OF INVENTION: SENSITIZATION OF CELLS TO CYTOTOXIC AGENTS USING

TITLE OF INVENTION: OLIGONUCLEOTIDES DIRECTED TO NUCLEOTIDE EXCISION REPAIR

; TITLE OF INVENTION: OR TRANSCRIPTION COUPLED REPAIR GENES
; FILE REFERENCE: HYZ-075US2 (475.08.514)
; CURRENT APPLICATION NUMBER: US/09/825,489
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
US-09-825-489-3

Query Match 100.0%; Score 20; DB 11; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
Db 1 GGTCCATCTCATGTTGATG 20

RESULT 2
US-09-867-701-7544
; Sequence 7544, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Agilate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7544
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(486)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-7544

Query Match 100.0%; Score 20; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
Db 149 GGTCCATCTCATGTTGATG 168

RESULT 3
US-09-867-701-2117/c
; Sequence 2117, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Agilate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2117
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(506)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2117

Query Match 100.0%; Score 20; DB 9; Length 506;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
Db 246 GGTCCATCTCATGTTGATG 227

RESULT 4
US-09-867-701-2107
; Sequence 2107, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Agilate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2107
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(513)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2107

Query Match 100.0%; Score 20; DB 9; Length 513;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
Db 357 GGTCCATCTCATGTTGATG 376

RESULT 5
US-09-825-489-13/c
; Sequence 13, Application US/09825489
; Publication No. US20030232767A1
; GENERAL INFORMATION:
; APPLICANT: AGRAMAL, SUDHIR
; APPLICANT: KANDIMALLA, EKAMBAR R.
; APPLICANT: BREGMAN, DAVID B.
; APPLICANT: MANI, SRIDHAR
; APPLICANT: LU, YI
; TITLE OF INVENTION: SENSITIZATION OF CELLS TO CYTOTOXIC AGENTS USING
; TITLE OF INVENTION: OLIGONUCLEOTIDES DIRECTED TO NUCLEOTIDE EXCISION REPAIR
; FILE REFERENCE: HYZ-075US2 (475.08.514)
; CURRENT APPLICATION NUMBER: US/09/825,489
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1377

TYPE: DNA
ORGANISM: Homo sapiens
US-09-825-489-13

Query Match 100.0%; Score 20; DB 11; Length 1377;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
|||||
DB 769 GGTCCTACTCATGTTGATG 750

RESULT 6
US-10-103-313-116/c
Sequence 116, Application US/10103313
Publication No. US20030082758A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ207C1
CURRENT APPLICATION NUMBER: US/10/103,313
CURRENT FILING DATE: 2002-03-12
NUMBER OF SEQ ID NOS: 653
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 116
SOFTWARE: Patentin Ver. 2.0
LENGTH: 1407
TYPE: DNA
ORGANISM: Homo sapiens
US-10-103-313-116

Query Match 100.0%; Score 20; DB 15; Length 1407;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
|||||
DB 782 GGTCCTACTCATGTTGATG 763

RESULT 7
US-10-101-510-587
Sequence 587, Application US/10101510
Publication No. US20030148295A1
GENERAL INFORMATION:
APPLICANT: WANG, JACKSON
APPLICANT: WANG, YIXIN
TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
FILE REFERENCE: 15117.0012
CURRENT APPLICATION NUMBER: US/10/101,510
CURRENT FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: 60/276,947
PRIOR FILING DATE: 2001-03-20
NUMBER OF SEQ ID NOS: 805
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 587
LENGTH: 4670
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (1781)..(1894)
OTHER INFORMATION: a, t, c, g, other or unknown
FEATURE:
NAME/KEY: modified_base
LOCATION: (2658)..(3015)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-587

Query Match 100.0%; Score 20; DB 15; Length 4670;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
|||||
DB 695 GGTCCTACTCATGTTGATG 714

RESULT 8
US-10-027-632-257254/c
Sequence 257254, Application US/10027632
Publication No. US20020196371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 257254
LENGTH: 1526
TYPE: DNA
ORGANISM: Human
US-10-027-632-257254

Query Match 85.0%; Score 17; DB 13; Length 1526;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CCATACATCATGTTGATG 20
|||||
DB 926 CCATACATCATGTTGATG 910

RESULT 9
US-10-027-632-257254/c
Sequence 257254, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 257254
LENGTH: 1526
TYPE: DNA
ORGANISM: Human
US-10-027-632-257254

Query Match
Best Local Similarity 85.0%; Score 17; DB 16; Length 1526;
Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTTGATG 20
|||||
Db 926 CCATCTCATGTTGATG 910

RESULT 10
US-10-158-844-40/C
Sequence 40, Application US/10158844
Publication No. US20040029118A1
GENERAL INFORMATION:

APPLICANT: Kunsch et al.
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude Pentium 3
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/158,844
FILING DATE: 03-Jun-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/961,527
FILING DATE: 1997-10-30
APPLICATION NUMBER: US 60/029,960
FILING DATE: 1996-10-31

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB340P1D1

SEQUENCE CHARACTERISTICS:

LENGTH: 14273 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 40:

US-10-158-844-40

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 13;
Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20
|||||
Db 13385 GGTCCATCTCATGTTGATG 13366

RESULT 11
US-10-027-632-253417/C
Sequence 253417, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 253417
LENGTH: 1049
TYPE: DNA
ORGANISM: Human
US-10-027-632-253417

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 13;
Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTTGATG 20
|||||
Db 71 TCCATCTCATGTTGATG 54

RESULT 12

US-10-027-632-253417/C
Sequence 253417, Application US/10027632
Publication No. US20030204075A9

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 253417

LENGTH: 1049

TYPE: DNA

ORGANISM: Human

US-10-027-632-253417

Query Match
Best Local Similarity 82.0%; Score 16.4; DB 16;
Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGATG 20
|||||
DB 71 TCCATTCATCATGTTGATG 54

RESULT 13
US-10-085-783A-40734
; Sequence 40734, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40734
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-40734

Query Match 79.0%; Score 15.8; DB 13; Length 238;
Best Local Similarity 89.5%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20
|||||
DB 77 GTCCATATCAATCAAGTTGATG 95

RESULT 14
US-10-242-535A-40734
; Sequence 40734, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40734
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-40734

Query Match 79.0%; Score 15.8; DB 16; Length 238;
Best Local Similarity 89.5%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20
|||||

DB 77 GTCCATATCAATCAAGTTGATG 95

RESULT 15
US-09-732-627A-2949/C
; Sequence 2949, Application US/09732627A
; Publication No. US20040123338A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(51770)B
; CURRENT APPLICATION NUMBER: US/09/732,627A
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4930
; SEQ ID NO 2949
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3493-010-P1-M1-E9
US-09-732-627A-2949

Query Match 79.0%; Score 15.8; DB 12; Length 296;
Best Local Similarity 89.5%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20
|||||
DB 29 GTCCATACATCATGTTGATG 11

RESULT 16
US-10-060-036-3118/C
; Sequence 3118, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3118
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 111..284
; OTHER INFORMATION: n = A,T,C or G
US-10-060-036-3118

Query Match 79.0%; Score 15.8; DB 15; Length 431;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20
|||||
DB 194 GTCCATATCAATCAAGTTGATG 176

RESULT 17
US-09-918-995-12216/C
; Sequence 12216, Application US/09918995
; Publication No. US20030073623A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
/ FROM VARIOUS CDNA LIBRARIES
/ FILE REFERENCE: 20411-756
/ CURRENT APPLICATION NUMBER: US/09/918,995
/ PRIOR FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: US/09/235,076
/ PRIOR FILING DATE: 1998-01-20
/ NUMBER OF SEQ ID NOS: 38054
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 12216
/ LENGTH: 462
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(462)
/ OTHER INFORMATION: n = A,T,C or G
US-09-918-995-12216

Query Match          79.0%; Score 15.8; DB 10; Length 462;
Best Local Similarity 89.5%; Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GTCCATCATCATGTTGATG 20
Db      446  GTCCCTCTCATGTTGATG 428

RESULT 18
US-10-085-783A-45577/C
/ Sequence 45577, Application US/10085783A
/ Publication No. US20040037841A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2002
/ CURRENT APPLICATION NUMBER: US/10/085,783A
/ CURRENT FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 45577
/ LENGTH: 540
/ TYPE: DNA
/ ORGANISM: Human
US-10-085-783A-45577

Query Match          79.0%; Score 15.8; DB 13; Length 540;
Best Local Similarity 89.5%; Pred. No. 6.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GGTCATCATCATGTTGAT 19
Db      233  GGTCACATCATGTTGAT 215

RESULT 19
US-10-242-535A-45577/C
/ Sequence 45577, Application US/10242535A
/ Publication No. US20040013663A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2005
```

```
/ CURRENT APPLICATION NUMBER: US/10/242,535A
/ CURRENT FILING DATE: 2002-09-12
/ PRIOR APPLICATION NUMBER: US 10/085,783
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 45577
/ LENGTH: 540
/ TYPE: DNA
/ ORGANISM: Human
US-10-242-535A-45577

Query Match          79.0%; Score 15.8; DB 16; Length 540;
Best Local Similarity 89.5%; Pred. No. 6.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GTCCATCATCATGTTGAT 19
Db      233  GTCCACATCATGTTGAT 215

RESULT 20
US-09-871-161-62
/ Sequence 62, Application US/09871161
/ Publication No. US20030097666A1
/ GENERAL INFORMATION:
/ APPLICANT: ENDBERG, WILSON O., ET AL.
/ TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
/ FILE REFERENCE: CCDNA-260XX
/ CURRENT APPLICATION NUMBER: US/09/871,161
/ CURRENT FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 09/328,111
/ PRIOR FILING DATE: 1999-06-08
/ PRIOR APPLICATION NUMBER: 60/117,393
/ PRIOR FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: 60/096,639
/ PRIOR FILING DATE: 1998-08-31
/ NUMBER OF SEQ ID NOS: 544
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 62
/ LENGTH: 614
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(614)
/ OTHER INFORMATION: n = A,T,C or G
US-09-871-161-62

Query Match          79.0%; Score 15.8; DB 10; Length 614;
Best Local Similarity 89.5%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GTCCATCATCATGTTGATG 20
Db      466  GTCCATATCAAGTTGATG 484

RESULT 21
US-10-641-643-1225/C
/ Sequence 1225, Application US/10641643
/ Publication No. US20040077003A1
/ GENERAL INFORMATION:
/ APPLICANT: Cocks, Benjamin G.
/ APPLICANT: Susan G. Stuart
/ APPLICANT: Jeffrey J. Seilhamer
```

```
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
/ GENE EXPRESSION
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/641,643
/ FILING DATE: 14-Aug-2003
/ CLASSIFICATION: <Unknown>
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: <Unknown>
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1225:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 686 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: G246741
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1225 :
US-10-641-643-1225

Query Match          79.0%; Score 15.8; DB 17; Length 686;
Best Local Similarity 89.5%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GGTCCATCTCATGTTGAT 19
DB      628  GGTCCATCTCATGTTGAT 610

RESULT 22
US-10-424-599-20394
/ Sequence 20394, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 20394
/ LENGTH: 765
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_118420C.1
US-10-424-599-20394
```

```
Query Match          79.0%; Score 15.8; DB 13; Length 765;
Best Local Similarity 89.5%; Pred. No. 6.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GTCCATCTCATGTTGATG 20
DB      484  GTCCATCTCATGTTGATG 502

RESULT 23
US-10-437-963-64475
/ Sequence 64475, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 64475
/ LENGTH: 789
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)-(789)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_65615C.1
US-10-437-963-64475

Query Match          79.0%; Score 15.8; DB 17; Length 789;
Best Local Similarity 89.5%; Pred. No. 6.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GTCCATCTCATGTTGATG 20
DB      442  GTCCATCTCATGTTGATG 460

RESULT 24
US-10-437-963-7531
/ Sequence 7531, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 7531
/ LENGTH: 1329
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
```

OTHER INFORMATION: Clone ID: PAT_MRT4530_14117C.1
US-10-437-963-7531

Query Match 79.0%; Score 15.8; DB 17; Length 1329;
Best Local Similarity 89.5%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
DB 355 GTCCATCTCATGTTGATG 373

RESULT 25
US-10-041-018-111

Sequence 111, Application US/10041018
Publication No. US20040072323A1
GENERAL INFORMATION:
APPLICANT: Matsuda, Seichi P.T.
APPLICANT: Hart, Elizabeth A.
TITLE OF INVENTION: Diterpene-Producing Unicellular Organism
FILE REFERENCE: P02080US1/10025547
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: US 60/259880
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 413
SOFTWARE: PatentIn version 3.1
SEQ ID NO 111
LENGTH: 3021
TYPE: DNA
ORGANISM: Dendroctonus jeffreyi
US-10-041-018-111

Query Match 79.0%; Score 15.8; DB 12; Length 3021;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
DB 2142 GTCCATCTCATGTTGATG 2160

RESULT 26
US-10-041-018-112

Sequence 112, Application US/10041018
Publication No. US20040072323A1
GENERAL INFORMATION:
APPLICANT: Matsuda, Seichi P.T.
APPLICANT: Hart, Elizabeth A.
TITLE OF INVENTION: Diterpene-Producing Unicellular Organism
FILE REFERENCE: P02080US1/10025547
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: US 60/259880
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 413
SOFTWARE: PatentIn version 3.1
SEQ ID NO 112
LENGTH: 5874
TYPE: DNA
ORGANISM: Dendroctonus jeffreyi
US-10-041-018-112

Query Match 79.0%; Score 15.8; DB 12; Length 5874;
Best Local Similarity 89.5%; Pred. No. 9.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
DB 4943 GTCCATCTCATGTTGATG 4961

RESULT 27

US-10-087-192-460
Sequence 460, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 529452000122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 460
LENGTH: 101209
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(101209)
OTHER INFORMATION: n = A,T,C or G
US-10-087-192-460

Query Match 79.0%; Score 15.8; DB 13; Length 101209;
Best Local Similarity 89.5%; Pred. No. 1.5e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
DB 39900 GTCCATCTCATGTTGATG 39918

RESULT 28
US-10-027-632-79848/C

Sequence 79848, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 79848
LENGTH: 519
TYPE: DNA
ORGANISM: Human
US-10-027-632-79848

Query Match 77.0%; Score 15.4; DB 13; Length 519;
Best Local Similarity 94.1%; Pred. No. 9.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTTGATG 20
 |||||
 Db 253 CCATACCCATGTTGATG 237

RESULT 29
 US-10-027-632-79849/c

; Sequence 79849, Application US/10027632
 ; Publication No. US20020198371A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 79849

; LENGTH: 519

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-79849

Query Match 77.0%; Score 15.4; DB 13; Length 519;
 Best Local Similarity 94.1%; Pred. No. 9.6e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTTGATG 20
 |||||
 Db 253 CCATACCCATGTTGATG 237

RESULT 30
 US-10-027-632-79848/c

; Sequence 79848, Application US/10027632
 ; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 79848
 ; LENGTH: 519
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-79848

Query Match 77.0%; Score 15.4; DB 16; Length 519;
 Best Local Similarity 94.1%; Pred. No. 9.6e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTTGATG 20
 |||||
 Db 253 CCATACCCATGTTGATG 237

RESULT 31
 US-10-027-632-79849/c

; Sequence 79849, Application US/10027632
 ; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 79849

; LENGTH: 519

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-79849

Query Match 77.0%; Score 15.4; DB 16; Length 519;
 Best Local Similarity 94.1%; Pred. No. 9.6e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTTGATG 20
 |||||
 Db 253 CCATACCCATGTTGATG 237

RESULT 32
 US-09-764-847-497

; Sequence 497, Application US/09764847
 ; Patent No. US20020132767A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC009

; CURRENT APPLICATION NUMBER: US/09/764,847

; PRIOR FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2003

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 497

; LENGTH: 551

```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (398)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (477)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (478)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (479)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (532)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-847-497
```

```
Query Match          77.0%; Score 15.4; DB 9; Length 551;
Best Local Similarity 94.1%; Pred. No. 9.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      3  TCCATCTCATGTTGAT 19
         |||||
Db      261 TCCATCTCATGTTGAT 277
```

```
RESULT 33
US-10-092-154-497
/ Sequence 497, Application US/10092154
/ Publication No. US20030054375A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PC009C1
/ CURRENT APPLICATION NUMBER: US/10/092,154
/ CURRENT FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 2003
/ Prior Application removed - See File Wrapper or Palm
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 497
/ LENGTH: 551
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (398)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (477)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (478)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (479)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (532)
/ OTHER INFORMATION: n equals a,t,g, or c
US-10-092-154-497
```

```
Query Match          77.0%; Score 15.4; DB 15; Length 551;
Best Local Similarity 94.1%; Pred. No. 9.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      3  TCCATCTCATGTTGAT 19
         |||||
Db      261 TCCATCTCATGTTGAT 277
```

```
RESULT 34
US-10-027-632-66707/c
/ Sequence 66707, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 66707
/ LENGTH: 573
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-66707
```

```
Query Match          77.0%; Score 15.4; DB 13; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      3  TCCATCTCATGTTGAT 19
         |||||
Db      471 TCCATCTCATGTTGAT 455
```

```
RESULT 35
US-10-027-632-66708/c
/ Sequence 66708, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 66708
/ LENGTH: 573
/ TYPE: DNA
/ ORGANISM: Human
```

US-10-027-632-66708

Query Match 77.0%; Score 15.4; DB 13; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19
|||||
Db 471 TCCATCTCATGTGTAT 455

RESULT 36

US-10-027-632-311197/c
; Sequence 311197, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 311196
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-311196

Query Match 77.0%; Score 15.4; DB 13; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19
|||||
Db 471 TCCATCTCATGTGTAT 455

RESULT 37

US-10-027-632-311197/c
; Sequence 311197, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 311197
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-311197

Query Match 77.0%; Score 15.4; DB 13; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19
|||||
Db 471 TCCATCTCATGTGTAT 455

RESULT 38
US-10-027-632-66707/c
; Sequence 66707, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66707
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-66707

Query Match 77.0%; Score 15.4; DB 16; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19
|||||
Db 471 TCCATCTCATGTGTAT 455

RESULT 39
US-10-027-632-66708/c
; Sequence 66708, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129


```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66708
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-66708

Query Match          77.0%; Score 15.4; DB 16; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTGAT 19
    |||||
Db 471 TCCATCTCATGTGAT 455

RESULT 40
US-10-027-632-311196/c
; Sequence 311196, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 311196
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-311196

Query Match          77.0%; Score 15.4; DB 16; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTGAT 19
    |||||
Db 471 TCCATCTCATGTGAT 455
```

```

RESULT 41
US-10-027-632-311197/c
; Sequence 311197, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 311197
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-311197

Query Match          77.0%; Score 15.4; DB 16; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTGAT 19
    |||||
Db 471 TCCATCTCATGTGAT 455

RESULT 42
US-10-424-599-103050
; Sequence 103050, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 103050
; LENGTH: 715
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MKT3847_64071C.1
US-10-424-599-103050

Query Match          77.0%; Score 15.4; DB 13; Length 715;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTGATG 20
    |||||
Db 471 TCCATCTCATGTGAT 455
```

Db 368 CCATGCTCATGTTGATG 384

RESULT 43

```
US-10-424-599-23361/c
; Sequence 23361, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285694
; SEQ ID NO 23361
; LENGTH: 3767
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121097C.1
US-10-424-599-23361
```

Query Match 77.0%; Score 15.4; DB 13; Length 3767;
Best Local Similarity 94.1%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGA 18
Db 2838 GTCCATCTCATGTTGA 2822

RESULT 44

```
US-10-087-192-1807/c
; Sequence 1807, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1807
; LENGTH: 119414
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(119414)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1807
```

Query Match 77.0%; Score 15.4; DB 13; Length 119414;
Best Local Similarity 94.1%; Pred. No. 2.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGA 18
Db 11805 GTCCATCTCATGTTGA 11789

RESULT 45

US-09-754-853A-2/c
; Sequence 2, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:

```
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 2
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45163)..(45314),(45450)..(45509),(46941)..(48763),(48975)..(49573)
; OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-2
```

Query Match 77.0%; Score 15.4; DB 10; Length 335913;
Best Local Similarity 94.1%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATCTCATGTTGAT 19
Db 24565 TCCAACTCATGTTGAT 24549

RESULT 46

```
US-09-754-853A-3/c
; Sequence 3, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 3
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46798)..(48763),(48975)..(49573)
; OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-3
```

Query Match 77.0%; Score 15.4; DB 10; Length 335913;
Best Local Similarity 94.1%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATCTCATGTTGAT 19
Db 24565 TCCAACTCATGTTGAT 24549

RESULT 47

US-10-424-599-127062/c

```
; Sequence 127062, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 127062
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85743C.1
US-10-424-599-127062

Query Match          76.0%; Score 15.2; DB 13; Length 292;
Best Local Similarity 85.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GGTCATCTCATGTGTGATG 20
Db      37  GGTCATCTCATGTGTGATG 18

RESULT 48
US-08-781-986A-1137
; Sequence 1137, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-1137

Query Match          76.0%; Score 15.2; DB 8; Length 330;
```

```
Best Local Similarity 85.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GGTCATCTCATGTGTGATG 20
Db      284  GGACCAACTCATGTGTGATG 303

RESULT 49
US-10-329-624-1137
; Sequence 1137, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; APPLICANT: Gil H. Choi
; APPLICANT: Patrick S. Dillon
; APPLICANT: Craig A. Rosen
; APPLICANT: Steven C. Barash
; APPLICANT: Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1137:
US-10-329-624-1137

Query Match          76.0%; Score 15.2; DB 13; Length 330;
Best Local Similarity 85.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GGTCATCTCATGTGTGATG 20
Db      284  GGACCAACTCATGTGTGATG 303

RESULT 50
US-10-424-599-130364
; Sequence 130364, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
```

```

1  APPLICANT:  La Rosa Thomas J
2  APPLICANT:  Kovalic David K
3  APPLICANT:  Zhou yihua
4  APPLICANT:  Cao Yongwei
5  TITLE OF INVENTION:  Soy Nucleic Acid Molecules and Other Molecules Associated With
6  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
7  FILE REFERENCE:  38-21(53223)B
8  CURRENT APPLICATION NUMBER:  US/10/424,599
9  CURRENT FILING DATE:  2003-04-28
10 NUMBER OF SEQ ID NOS:  285684
11
12 SEQ ID NO 130364
13
14 LENGTH:  358
15
16 TYPE:  DNA
17
18 ORGANISM:  Glycine max
19
20 FEATURE:
21
22 OTHER INFORMATION:  Clone ID:  PAT_MRT3847_88727C.1
23
24 US-10-424-599-130364

```

Query Match	Similarity	76.0%	Score	15.2	DB	length	358	
Best local	Similarity	85.0%	Pred. No.	1	le	03		
Matches	17; Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
Qy	1	GGTCCTACTCATGTGGATG	20					
Db	97	GGCTCAATATCTTGTGGATG	116					

Search completed: August 17, 2004, 16:56:43
Job time : 396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 13:20:27 ; Search time 1548 Seconds
(without alignments)
559.988 Million cell updates/sec

Title: US-09-825-489-3

Perfect score: 20

Sequence: 1 ggtccatctctgtgtgatg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 800 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_scs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_scs:*

28: em_un:*

29: em_vl:*

30: em_hlg_hum:*

31: em_hlg_inv:*

32: em_hlg_other:*

33: em_hlg_mus:*

34: em_hlg_pin:*

35: em_hlg_rnd:*

36: em_hlg_rnd:*

37: em_hlg_vrt:*

38: em_hlg_vrt:*

39: em_hgo_hum:*

40: em_hgo_mus:*

41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	AX282880	AX282880 Sequence
2	20	100.0	809	HSXPAC5	U10347 Human Xerod
3	20	100.0	822	AR183106	AR183106 Sequence
4	20	100.0	822	AR203243	AR203243 Sequence
5	20	100.0	822	AX277000	AX277000 Sequence
6	20	100.0	1377	AX282890	AX282890 Sequence
7	20	100.0	1377	HMXKPA6	D14533 Homo sapien
8	20	100.0	1439	BC014965	BC014965 Homo sapi
9	20	100.0	24994	AF503166	AF503166 Homo sapi
10	20	100.0	111345	AL445531	AL445531 Human DNA
11	20	100.0	187079	AL442130	AL442130 Homo sapi
12	18.4	92.0	1095	XELXPACB2	D31895 Xenopus lae
13	18.4	92.0	1277	XELXPACB	D31894 Xenopus lae
14	18.4	92.0	157653	AL732555	AL732555 Mouse DNA
15	18.4	92.0	199263	AC109694	AC109694 Rattus no
16	18.4	92.0	227524	AC126180	AC126180 Rattus no
17	17.4	87.0	158826	AP003607	AP003607 Oryza sat
18	17.4	87.0	197674	AP004367	AP004367 Oryza sat
19	17.4	87.0	198448	AC115118	AC115118 Mus muscu
20	17	85.0	189134	HDAC004682	AC004682 Homo sapi
21	17	85.0	190856	AC009160	AC009160 Homo sapi
22	17	85.0	253297	AC134076	AC134076 Rattus no
23	17	85.0	263730	AC094314	AC094314 Rattus no
24	16.8	84.0	294	MXKPA6	X74350 M. musculus
25	16.8	84.0	576	AF354052	AF354052 Mus muscu
26	16.8	84.0	938	MXKPA7	X74351 M. musculus
27	16.8	84.0	1029	AX566974	AX566974 Sequence
28	16.8	84.0	2138	AK115474	AK115474 Ciona int
29	16.8	84.0	11344	AE007333	AE007333 Streptoco
30	16.8	84.0	14273	AR218808	AR218808 Sequence
31	16.8	84.0	14273	BD003720	BD003720 Polynucle
32	16.8	84.0	90169	HS398A12	AL021332 Human DNA
33	16.8	84.0	110000	AC095009-1	Continuation (2 of
34	16.8	84.0	121113	BX537312	BX537312 Danio rer
35	16.8	84.0	121852	AP003965	AP003965 Homo sapi
36	16.8	84.0	127247	HS326133	Z82170 Human DNA
37	16.8	84.0	129300	AP003886	AP003886 Oryza sat
38	16.8	84.0	148416	AC002089	AC002089 Homo sapi
39	16.8	84.0	152635	AC011144	AC011144 Homo sapi
40	16.8	84.0	155646	AC146044	AC146044 Pan trogl
41	16.8	84.0	157200	AP005233	AP005233 Homo sapi
42	16.8	84.0	164572	OSJN00128	AL607001 Oryza sat
43	16.8	84.0	175936	SPNEU108	AL449930 Streptoco
44	16.8	84.0	186184	AC139894	AC139894 Rattus no
45	16.8	84.0	209706	AC119611	AC119611 Rattus no
46	16.8	84.0	213381	AC129598	AC129598 Mus muscu
47	16.8	84.0	216911	AC107756	AC107756 Mus muscu
48	16.8	84.0	225182	AC107742	AC107742 Mus muscu
49	16.8	84.0	300340	AB016936	AB016936 Bacteroid
50	16.8	84.0	349980	AX571760	AX571760 Sequence
51	16.4	82.0	3573	LLJ000883	AJ000883 Lactococc
52	16.4	82.0	6765	AX586945	AX586945 Sequence
53	16.4	82.0	7568	AF092042	AF092042 Lactococc
54	16.4	82.0	10240	AB006382	AB006382 Lactococc
55	16.4	82.0	15246	AX586923	AX586923 Sequence
56	16.4	82.0	22668	AF536818	AF536818 Homo sapi
57	16.4	82.0	49321	AF275366	AF275366 Mus muscu
58	16.4	82.0	56758	AC101049	AC101049 Mus muscu
59	16.4	82.0	63489	AP002864	AP002864 Oryza sat
60	16.4	82.0	92487	AC068280	AC068280 Homo sapi
61	16.4	82.0	93388	AC105432	AC105432 Magnapor
62	16.4	82.0	93876	AL390240	AL390240 Human DNA
63	16.4	82.0	110000	AC114905-1	Continuation (2 of
64	16.4	82.0	115768	AP000712	AP000712 Homo sapi
65	16.4	82.0	137411	AC124678	AC124678 Mus muscu

C 66	16.4	82.0	137569	2	AC084093	AC084093 Homo sapi	C 139	15.8	79.0	1804	10	NMST6SIAD	X94000 M. musculus
C 67	16.4	82.0	140730	9	AP001960	AP001960 Homo sapi	C 140	15.8	79.0	1844	8	AK103896	AK103896 Oryza sat
C 68	16.4	82.0	141272	9	AP001976	AP001976 Homo sapi	C 141	15.8	79.0	2018	3	AY113567	AY113567 Drosophill
C 69	16.4	82.0	144383	8	AC092780	AC092780 Oryza sat	C 142	15.8	79.0	2282	8	BC000148	BC000148 Homo sapi
C 70	16.4	82.0	148501	9	BX284929	BX284929 Human DNA	C 143	15.8	79.0	2285	8	AK118438	AK118438 Arabidope
C 71	16.4	82.0	150394	2	BX005147	BX005147 Dantio rer	C 144	15.8	79.0	2302	9	BC003686	BC003686 Homo sapi
C 72	16.4	82.0	150858	8	AC091246	AC091246 Oryza sat	C 145	15.8	79.0	2324	9	BC022890	BC022890 Homo sapi
C 73	16.4	82.0	152858	8	AC093148	AC093148 Homo sapi	C 146	15.8	79.0	2712	7	BR034036	BR034036 Bacteriopho
C 74	16.4	82.0	154937	2	AP002382	AP002382 Homo sapi	C 147	15.8	79.0	2963	3	AF435855	AF435855 Plasmodiu
C 75	16.4	82.0	154918	2	AP004866	AP004866 Oryza sat	C 148	15.8	79.0	3021	3	AF159136	AF159136 Dendrocto
C 76	16.4	82.0	159736	10	AC122292	AC122292 Mus muscu	C 149	15.8	79.0	4039	14	HSB152A	101099 Bovine herd
C 77	16.4	82.0	161348	10	AL645352	AL645352 Mouse DNA	C 150	15.8	79.0	5874	3	AF159137S2	AF159138 Dendrocto
C 78	16.4	82.0	164502	2	AC026960	AC026960 Homo sapi	C 151	15.8	79.0	10671	1	AE007842	AE007842 Clostrid
C 79	16.4	82.0	167084	8	AP002999	AP002999 Homo sapi	C 152	15.8	79.0	11034	14	AF104985	AF104985 Hixame rh
C 80	16.4	82.0	168552	8	AP004891	AP004891 Oryza sat	C 153	15.8	79.0	15033	2	AC012888	AC012888 Drosophill
C 81	16.4	82.0	168875	9	AC114794	AC114794 Homo sapi	C 154	15.8	79.0	18900	8	AY032722	AY032722 Homo sapi
C 82	16.4	82.0	169232	2	AC024299	AC024299 Homo sapi	C 155	15.8	79.0	20149	8	CNS086C7M	AL731737 Oryza sat
C 83	16.4	82.0	170212	2	AC058822	AC058822 Homo sapi	C 156	15.8	79.0	24006	2	AC014743	AC014743 Drosophill
C 84	16.4	82.0	177878	2	BX469895	BX469895 Dantio rer	C 157	15.8	79.0	29545	2	AC112696	AC112696 Homo sapi
C 85	16.4	82.0	178367	2	AC015614	AC015614 Homo sapi	C 158	15.8	79.0	30583	2	AC017314	AC017314 Drosophill
C 86	16.4	82.0	181123	2	AC104064	AC104064 Homo sapi	C 159	15.8	79.0	35560	2	AC127494	AC127494 Homo sapi
C 87	16.4	82.0	182603	9	AF127019	AF127019 Homo sapi	C 160	15.8	79.0	50374	2	AC015170	AC015170 Drosophill
C 88	16.4	82.0	185758	10	AC102173	AC102173 Mus muscu	C 161	15.8	79.0	50607	2	AC090598	AC090598 Homo sapi
C 89	16.4	82.0	187131	2	AP001535	AP001535 Homo sapi	C 162	15.8	79.0	57287	2	AC008283	AC008283 Homo sapi
C 90	16.4	82.0	188927	10	AC120136	AC120136 Mus muscu	C 163	15.8	79.0	61173	1	AY372454	AY372454 Unculture
C 91	16.4	82.0	189949	2	AC110260	AC110260 Mus muscu	C 164	15.8	79.0	61532	2	AC144396	AC144396 Rattus no
C 92	16.4	82.0	192230	2	AC027065	AC027065 Homo sapi	C 165	15.8	79.0	61532	2	AC144396	AC144396 Rattus no
C 93	16.4	82.0	194518	9	AC113147	AC113147 Homo sapi	C 166	15.8	79.0	71396	2	AC100674	AC100674 Mus muscu
C 94	16.4	82.0	195410	9	AC096639	AC096639 Homo sapi	C 167	15.8	79.0	72387	2	AC005811	AC005811 Drosophill
C 95	16.4	82.0	199027	2	AC026611	AC026611 Homo sapi	C 168	15.8	79.0	73148	2	AC048349	AC048349 Homo sapi
C 96	16.4	82.0	204860	9	AC083902	AC083902 Homo sapi	C 169	15.8	79.0	73170	9	AL445068	AL445068 Human DNA
C 97	16.4	82.0	205594	9	AP002802	AP002802 Homo sapi	C 170	15.8	79.0	75499	9	HS0965G21	AL121772 Human DNA
C 98	16.4	82.0	207303	2	AL645562	AL645562 Homo sapi	C 171	15.8	79.0	86825	2	AC131082	AC131082 Mus muscu
C 99	16.4	82.0	211844	2	AP0048370	AP0048370 Homo sapi	C 172	15.8	79.0	90624	10	AL731671	AL731671 Mouse DNA
C 100	16.4	82.0	217393	9	AP001122	AP001122 Homo sapi	C 173	15.8	79.0	90882	2	AL139282	AL139282 Human DNA
C 101	16.4	82.0	226637	10	AL731843	AL731843 Mouse DNA	C 174	15.8	79.0	92059	2	AC128823_3	AC128823_3
C 102	16.4	82.0	234976	2	AC118698	AC118698 Mus muscu	C 175	15.8	79.0	92652	8	AC024261	AC024261 Arabidops
C 103	16.4	82.0	237845	2	AC112433	AC112433 Rattus no	C 176	15.8	79.0	93296	9	AL136363	AL136363 Human DNA
C 104	16.4	82.0	243479	2	BX537152	BX537152 Dantio rer	C 177	15.8	79.0	99791	8	CNS07Y00	AL713900 Oryza sat
C 105	16.4	82.0	244441	2	AC098180	AC098180 Rattus no	C 178	15.8	79.0	102092	2	AC136053	AC136053 Rattus no
C 106	16.4	82.0	249412	2	AC095097	AC095097 Rattus no	C 179	15.8	79.0	108873	14	AF136573	AF136573 Bovine he
C 107	16.4	82.0	254004	2	AC129469	AC129469 Rattus no	C 180	15.8	79.0	110000	2	AC098193_4	AC098193_4
C 108	16.4	82.0	257078	2	AC118300	AC118300 Rattus no	C 181	15.8	79.0	110000	2	AC122076_2	AC122076_2
C 109	16.4	82.0	257784	2	AC1118627	AC1118627 Mus muscu	C 182	15.8	79.0	110000	2	BX294172_1	BX294172_1
C 110	16.4	82.0	260480	2	AC141647	AC141647 Mus muscu	C 183	15.8	79.0	112405	9	AC072023	AC072023 Homo sapi
C 111	16.4	82.0	260657	2	AC094371	AC094371 Mus muscu	C 184	15.8	79.0	112528	2	AC140068	AC140068 Medicago
C 112	16.4	82.0	300250	10	AF125314	AF125314 Mus muscu	C 185	15.8	79.0	119974	9	AC018763	AC018763 Homo sapi
C 113	16.4	82.0	51112	8	AF492377	AF492377 Aegilops	C 186	15.8	79.0	122079	9	AC009401	AC009401 Homo sapi
C 114	16.4	82.0	81175	9	AL445245	AL445245 Human DNA	C 187	15.8	79.0	122815	8	AC135563	AC135563 Oryza sat
C 115	16.4	82.0	81705	8	AP006384	AP006384 Lotus cor	C 188	15.8	79.0	125933	9	AL732586	AL732586 Human DNA
C 116	16.4	82.0	97865	8	AP006093	AP006093 Lotus cor	C 189	15.8	79.0	128758	1	AF279106	AF279106 Unculture
C 117	16.4	82.0	119684	9	AC008849	AC008849 Homo sapi	C 190	15.8	79.0	131133	8	AC146683	AC146683 Medicago
C 118	16.4	82.0	134319	9	AC113405	AC113405 Homo sapi	C 191	15.8	79.0	131441	8	CNS086CDA	AL954871 Oryza sat
C 119	16.4	82.0	148415	2	AC113381	AC113381 Homo sapi	C 192	15.8	79.0	132221	2	AC102001	AC102001 Mus muscu
C 120	16.4	82.0	155895	2	AC102166	AC102166 Mus muscu	C 193	15.8	79.0	138129	2	AC011036	AC011036 Homo sapi
C 121	16.4	82.0	160703	9	AC022104	AC022104 Homo sapi	C 194	15.8	79.0	139788	2	AC024118	AC024118 Homo sapi
C 122	16.4	82.0	174101	2	AC119206	AC119206 Mus muscu	C 195	15.8	79.0	142347	10	AC112545	AC112545 Mus muscu
C 123	16.4	82.0	179424	2	AC027222	AC027222 Homo sapi	C 196	15.8	79.0	142712	9	AC112175	AC112175 Homo sapi
C 124	16.4	82.0	181792	9	AC098822	AC098822 Homo sapi	C 197	15.8	79.0	143451	2	AC021295	AC021295 Homo sapi
C 125	16.4	82.0	225164	2	AC104203	AC104203 Mus muscu	C 198	15.8	79.0	144357	2	AC144685	AC144685 Rattus no
C 126	16.4	82.0	228632	2	AC114202	AC114202 Rattus no	C 199	15.8	79.0	145247	10	AC138173	AC138173 Mus muscu
C 127	16.4	82.0	238919	2	AC139513	AC139513 Mus muscu	C 200	15.8	79.0	146128	9	AC026358	AC026358 Homo sapi
C 128	16.4	82.0	239893	2	AC107521	AC107521 Rattus no	C 201	15.8	79.0	147376	10	AL607039	AL607039 Mouse DNA
C 129	15.8	79.0	250	11	DM15812T	Z31891 D. melanoga	C 202	15.8	79.0	149546	9	AC087256	AC087256 Homo sapi
C 130	15.8	79.0	614	6	BD229387	BD229387 Human gen	C 203	15.8	79.0	149819	2	AC121091	AC121091 Mus muscu
C 131	15.8	79.0	666	6	AK380680	AK380680 Sequence	C 204	15.8	79.0	149960	5	AC144706	AC144706 Dantio rer
C 132	15.8	79.0	666	6	S87068S9	S87068 CD8 beta 1=	C 205	15.8	79.0	150106	2	AC144913	AC144913 Mus muscu
C 133	15.8	79.0	801	6	AR375131	AR375131 Sequence	C 206	15.8	79.0	150105	8	CNS0639E	AL732639 Oryza sat
C 134	15.8	79.0	936	6	AK060174	AK060174 Oryza sat	C 207	15.8	79.0	151547	8	AP003312	AP003312 Oryza sat
C 135	15.8	79.0	1176	6	AR376767	AR376767 Sequence	C 208	15.8	79.0	151862	10	AC102383	AC102383 Mus muscu
C 136	15.8	79.0	1480	8	AK069322	AK069322 Oryza sat	C 209	15.8	79.0	152220	8	OSJN00166	AL662967 Oryza sat
C 137	15.8	79.0	1698	8	AK102189	AK102189 Oryza sat	C 210	15.8	79.0	152313	9	AL591594	AL591594 Human DNA
C 138	15.8	79.0	1796	8	AK071708	AK071708 Oryza sat	C 211	15.8	79.0	153203	2	AL359812	AL359812 Homo sapi

C 212	15.8	79.0	153307	10	AL845441	285	15.8	79.0	191814	10	AC130541
C 213	15.8	79.0	154369	9	AC004803	C 286	15.8	79.0	191856	9	AL355980
C 214	15.8	79.0	155713	5	BX000347	C 287	15.8	79.0	193391	2	AC026824
C 215	15.8	79.0	156225	2	AC115007	C 288	15.8	79.0	194224	5	AL954868
C 216	15.8	79.0	157310	2	RM187K23	C 289	15.8	79.0	194510	2	AC092956
C 217	15.8	79.0	157428	10	AL606832	C 290	15.8	79.0	195269	2	AC134458
C 218	15.8	79.0	158907	9	AC011890	C 291	15.8	79.0	195546	2	AC121701
C 219	15.8	79.0	159235	9	AL356954	C 292	15.8	79.0	195350	2	BX248319
C 220	15.8	79.0	159305	9	AL355384	C 293	15.8	79.0	196817	2	AC019209
C 221	15.8	79.0	159919	9	AL359194	C 294	15.8	79.0	196893	2	BX323059
C 222	15.8	79.0	161067	9	AL353638	C 295	15.8	79.0	199289	10	AL772275
C 223	15.8	79.0	161317	2	AL391873	C 296	15.8	79.0	199894	2	AC115053
C 224	15.8	79.0	161582	2	AC110650	C 297	15.8	79.0	200425	10	AC114549
C 225	15.8	79.0	163284	2	AL137000	C 298	15.8	79.0	200634	2	AC016418
C 226	15.8	79.0	163314	2	AL590557	C 299	15.8	79.0	202295	2	AC109142
C 227	15.8	79.0	163475	3	AC093099	C 300	15.8	79.0	203029	2	BX855594
C 228	15.8	79.0	164193	3	AC006494	C 301	15.8	79.0	203282	10	AC116674
C 229	15.8	79.0	165056	2	AC022412	C 302	15.8	79.0	204235	2	AC117423
C 230	15.8	79.0	165328	2	AC022185	C 303	15.8	79.0	204249	2	AC133268
C 231	15.8	79.0	165589	2	AC011052	C 304	15.8	79.0	204445	2	AC126512
C 232	15.8	79.0	165638	2	AC103514	C 305	15.8	79.0	205616	2	CNS05TDE
C 233	15.8	79.0	165964	2	AC128313	C 306	15.8	79.0	206431	2	AL935191
C 234	15.8	79.0	166185	5	BX649507	C 307	15.8	79.0	206438	5	AL935268
C 235	15.8	79.0	166652	2	BX537330	C 308	15.8	79.0	206784	9	HS9317
C 236	15.8	79.0	167560	7	AY303349	C 309	15.8	79.0	206860	2	AC102884
C 237	15.8	79.0	168062	9	AL353768	C 310	15.8	79.0	208367	2	AC141868
C 238	15.8	79.0	168227	9	AC021868	C 311	15.8	79.0	208561	9	AC023206
C 239	15.8	79.0	168599	10	AL935336	C 312	15.8	79.0	209109	9	AC116025
C 240	15.8	79.0	168909	2	AC118862	C 313	15.8	79.0	209643	10	AC080018
C 241	15.8	79.0	169158	2	AC121376	C 314	15.8	79.0	210074	9	AC010837
C 242	15.8	79.0	171004	2	AC006286	C 315	15.8	79.0	210332	10	AL808131
C 243	15.8	79.0	171209	4	AC124905	C 316	15.8	79.0	210537	2	AC109840
C 244	15.8	79.0	171454	3	AC007137	C 317	15.8	79.0	211509	2	CNS08CBD
C 245	15.8	79.0	171520	2	BX470069	C 318	15.8	79.0	212475	2	AC093021
C 246	15.8	79.0	171710	2	AC013388	C 319	15.8	79.0	213045	2	AC120846
C 247	15.8	79.0	171868	3	AC010118	C 320	15.8	79.0	214505	2	AC099590
C 248	15.8	79.0	171970	2	AL590454	C 321	15.8	79.0	214838	9	AC012501
C 249	15.8	79.0	171974	9	AC092042	C 322	15.8	79.0	215294	2	AC117585
C 250	15.8	79.0	172123	9	AC060809	C 323	15.8	79.0	215615	2	AC098139
C 251	15.8	79.0	172479	3	AC006590	C 324	15.8	79.0	219266	2	AC130080
C 252	15.8	79.0	172695	9	AC068303	C 325	15.8	79.0	220367	2	AC087133
C 253	15.8	79.0	173535	9	AC142305	C 326	15.8	79.0	220935	2	AC111951
C 254	15.8	79.0	173672	5	BX248086	C 327	15.8	79.0	224136	2	AC133260
C 255	15.8	79.0	173844	9	AC078842	C 328	15.8	79.0	224417	2	AC095923
C 256	15.8	79.0	173856	9	AC105383	C 329	15.8	79.0	224548	3	AE003739
C 257	15.8	79.0	175232	2	AC132922	C 330	15.8	79.0	226579	10	AC133525
C 258	15.8	79.0	175238	2	AC101745	C 331	15.8	79.0	227155	10	AL732593
C 259	15.8	79.0	176049	10	AL844198	C 332	15.8	79.0	230204	2	AC123996
C 260	15.8	79.0	176161	10	AC099578	C 333	15.8	79.0	231140	2	AC116231
C 261	15.8	79.0	176987	2	AC103356	C 334	15.8	79.0	231165	2	AC120231
C 262	15.8	79.0	177583	3	AC105264	C 335	15.8	79.0	233320	10	AC127314
C 263	15.8	79.0	178562	3	AC091954	C 336	15.8	79.0	233965	2	AC125739
C 264	15.8	79.0	178931	3	AC008197	C 337	15.8	79.0	234429	9	AF311103
C 265	15.8	79.0	179199	2	AC120237	C 338	15.8	79.0	234758	2	AC130851
C 266	15.8	79.0	180242	5	AC027620	C 339	15.8	79.0	234967	2	AC109380
C 267	15.8	79.0	180872	5	AL844185	C 340	15.8	79.0	235008	2	AC119443
C 268	15.8	79.0	181903	2	AC145923	C 341	15.8	79.0	235150	2	AC095650
C 269	15.8	79.0	181935	9	AC018728	C 342	15.8	79.0	236832	2	AC119783
C 270	15.8	79.0	182288	9	AC018362	C 343	15.8	79.0	237805	2	AC121047
C 271	15.8	79.0	182452	9	AL161731	C 344	15.8	79.0	238355	2	AC126306
C 272	15.8	79.0	186105	10	AC055777	C 345	15.8	79.0	239789	2	AC129342
C 273	15.8	79.0	186564	10	AC119816	C 346	15.8	79.0	240747	2	AC105839
C 274	15.8	79.0	187415	2	AC026352	C 347	15.8	79.0	241000	2	AC121369
C 275	15.8	79.0	188078	2	AC008045	C 348	15.8	79.0	241834	2	AC095798
C 276	15.8	79.0	188734	2	AC069443	C 349	15.8	79.0	241899	2	AC120228
C 277	15.8	79.0	188812	9	AL357055	C 350	15.8	79.0	242037	2	AC114200
C 278	15.8	79.0	189650	10	AC136091	C 351	15.8	79.0	242348	2	AC111361
C 279	15.8	79.0	189732	9	AC023034	C 352	15.8	79.0	242386	2	AC098053
C 280	15.8	79.0	190084	9	AC140950	C 353	15.8	79.0	242649	2	AC108648
C 281	15.8	79.0	190692	9	AC021379	C 354	15.8	79.0	243563	2	AC098187
C 282	15.8	79.0	190987	2	AC131883	C 355	15.8	79.0	244259	2	AC123226
C 283	15.8	79.0	191239	2	AP000883	C 356	15.8	79.0	249148	2	AC106405
C 284	15.8	79.0	191265	2	AC068996	C 357	15.8	79.0	249922	2	AC129001

C 358	15.8	79.0	250087	2	AC096931	Rattus no	431	15.4	77.0	144901	2	AC125276	AC135276 Mus muscu
359	15.8	79.0	250425	2	AC125667	Rattus no	C 432	15.4	77.0	145536	2	AC105166	AC105166 Mus muscu
360	15.8	79.0	251173	2	AC115487	Rattus no	C 433	15.4	77.0	149986	2	AC069256	AC069256 Homo sapi
361	15.8	79.0	251478	2	AC105849	Rattus no	C 434	15.4	77.0	150076	9	CNS01DPI	AL132126 Human chr
C 362	15.8	79.0	252602	2	AC091704	Mus muscu	C 435	15.4	77.0	153046	9	AL590009	AL590009 Human DNA
363	15.8	79.0	253185	2	AC132749	Rattus no	436	15.4	77.0	154234	3	AC010034	AC010034 Drosophill
364	15.8	79.0	257867	3	AC005557	Drosophill	C 437	15.4	77.0	157979	9	CNS05TD4	AL335592 Human chr
C 365	15.8	79.0	258416	3	AC098944	Rattus no	438	15.4	77.0	158375	2	AC080039	AC080039 Homo sapi
C 366	15.8	79.0	258710	2	AC137281	Rattus no	439	15.4	77.0	161602	9	AC022634	AC022634 Homo sapi
C 367	15.8	79.0	260027	3	AE003659	Drosophill	440	15.4	77.0	162835	10	AC121866	AC121866 Mus muscu
C 368	15.8	79.0	273995	3	AE003669	Drosophill	C 441	15.4	77.0	164932	10	AC121942	AC121942 Mus muscu
C 369	15.8	79.0	274047	2	AC134087	Rattus no	442	15.4	77.0	166020	2	AC136149	AC136149 Oryza sat
C 370	15.8	79.0	284127	2	AC128719	Rattus no	C 443	15.4	77.0	166138	2	AC084756	AC084756 Homo sapi
C 371	15.8	79.0	300134	1	AE001721	Geobacter	C 444	15.4	77.0	166606	2	AC020685	AC020685 Homo sapi
C 372	15.8	79.0	303191	3	AE003472	Drosophill	445	15.4	77.0	166766	2	AC119682	AC119682 Rattus no
373	15.8	79.0	330919	2	AC095296	Rattus no	446	15.4	77.0	168862	2	AC092577	AC092577 Padio anu
374	15.8	79.0	341560	2	AL596304	Homo sapi	447	15.4	77.0	168880	2	AC114315	AC114315 Homo sapi
375	15.4	77.0	633	6	AX142249	Sequence	448	15.4	77.0	169757	5	AL627256	AL627256 Zebrafish
376	15.4	77.0	675	6	AX143961	Sequence	C 449	15.4	77.0	170048	2	AC025240	AC025240 Homo sapi
C 377	15.4	77.0	754	10	BC048550	Mus muscu	C 450	15.4	77.0	170466	3	AC010012	AC010012 Drosophill
C 378	15.4	77.0	895	8	AF526257	Glycine m	C 451	15.4	77.0	170532	9	AL161728	AL161728 Human DNA
C 379	15.4	77.0	1420	8	AK026285	Homo sapi	C 452	15.4	77.0	171347	9	AC099776	AC099776 Homo sapi
C 380	15.4	77.0	1686	8	HYU310841	Hordeum v	C 453	15.4	77.0	172295	10	AL845460	AL845460 Mouse DNA
C 381	15.4	77.0	1934	8	AT147012	Medicago	C 454	15.4	77.0	173645	2	AC010770	AC010770 Homo sapi
C 382	15.4	77.0	2554	8	BD232696	Encodemen	C 455	15.4	77.0	173769	9	AL696819	AL696819 Homo sapi
C 383	15.4	77.0	2876	8	AK065951	Oryza sat	C 456	15.4	77.0	173962	2	AC024155	AC024155 Rattus no
384	15.4	77.0	2946	1	AF269698	Staphyloc	C 457	15.4	77.0	174002	2	AC123406	AC123406 Rattus no
385	15.4	77.0	2946	6	AX145016	Sequence	C 458	15.4	77.0	174393	3	AC011720	AC011720 Homo sapi
C 386	15.4	77.0	3021	6	AF269704	Staphyloc	C 459	15.4	77.0	174949	2	AC111342	AC111342 Rattus no
C 387	15.4	77.0	3021	6	AX145022	Sequence	C 460	15.4	77.0	175008	2	AC113300	AC113300 Mus muscu
C 388	15.4	77.0	3391	1	AF269717	Staphyloc	C 461	15.4	77.0	175876	2	AC123336	AC123336 Rattus no
389	15.4	77.0	3391	6	AX145035	Sequence	C 462	15.4	77.0	175946	9	AP002088	AP002088 Homo sapi
C 390	15.4	77.0	3725	10	BC042528	Mus muscu	C 463	15.4	77.0	176021	5	AC026562	AC026562 Homo sapi
C 391	15.4	77.0	3768	8	AF195029	Glycine m	C 464	15.4	77.0	176569	2	BX088540	BX088540 Zebrafish
C 392	15.4	77.0	11906	3	AF321003	Ostrinia	C 465	15.4	77.0	176865	2	AL513306	AL513306 Homo sapi
393	15.4	77.0	14536	3	AF467260	Mus muscu	C 466	15.4	77.0	176967	2	AC022742	AC022742 Homo sapi
C 394	15.4	77.0	39578	2	AC101097	Mus muscu	C 467	15.4	77.0	177246	2	AC114308	AC114308 Homo sapi
C 395	15.4	77.0	49976	2	AC025670	Homo sapi	C 468	15.4	77.0	177986	2	AC140745	AC140745 Rattus no
C 396	15.4	77.0	57144	2	AC100329	Mus muscu	C 469	15.4	77.0	178316	9	AC022690	AC022690 Homo sapi
C 397	15.4	77.0	63753	2	AL690089	Homo sapi	470	15.4	77.0	178364	10	AL645594	AL645594 Mouse DNA
C 398	15.4	77.0	64341	10	AL627323	Mouse DNA	C 471	15.4	77.0	178899	9	AL671860	AL671860 Mouse DNA
399	15.4	77.0	65225	2	AC079109	Homo sapi	C 472	15.4	77.0	179169	2	AC092759	AC092759 Padio anu
C 400	15.4	77.0	74282	8	AB025606	Arabidops	C 473	15.4	77.0	179463	2	AL596453	AL596453 Homo sapi
C 401	15.4	77.0	75076	8	AC004948	Homo sapi	474	15.4	77.0	179724	3	AP003094	AP003094 Homo sapi
C 402	15.4	77.0	79556	8	AC119567	Homo sapi	C 475	15.4	77.0	183297	2	AC102312	AC102312 Mus muscu
C 403	15.4	77.0	80009	3	AF226688	Bombyx mo	C 476	15.4	77.0	183547	2	AC009551	AC009551 Homo sapi
404	15.4	77.0	84653	2	BX663523	Gallus ga	477	15.4	77.0	184194	9	AC008952	AC008952 Homo sapi
C 405	15.4	77.0	89666	2	AC022403	Homo sapi	478	15.4	77.0	184688	9	AC092944	AC092944 Homo sapi
C 406	15.4	77.0	90015	9	AL513264	Human DNA	479	15.4	77.0	185024	2	AC027649	AC027649 Mus muscu
C 407	15.4	77.0	90599	8	AC130275	Medicago	C 480	15.4	77.0	185369	10	AL671860	AL671860 Mouse DNA
408	15.4	77.0	91073	9	AP001093	Homo sapi	481	15.4	77.0	186233	9	AC092329	AC092329 Homo sapi
C 409	15.4	77.0	103177	5	AP0000841	Homo sapi	482	15.4	77.0	186253	9	AC020892	AC020892 Homo sapi
C 410	15.4	77.0	106095	2	BX005474	Zebrafish	483	15.4	77.0	189317	2	AC024483	AC024483 Homo sapi
411	15.4	77.0	107093	2	BX663527	Gallus ga	484	15.4	77.0	189461	2	AC118247	AC118247 Mus muscu
C 412	15.4	77.0	110000	2	AC097394	Rattus no	C 485	15.4	77.0	190718	8	CNS08C7T	AL1731744 Oryza sat
C 413	15.4	77.0	110024	2	AC125752	Continuatio (2 of	486	15.4	77.0	193931	2	AC116486	AC116486 Mus muscu
414	15.4	77.0	113024	2	AL445260	Drosophill	C 487	15.4	77.0	197288	2	AC131584	AC131584 Homo sapi
C 415	15.4	77.0	113311	2	AC019822	Drosophill	C 488	15.4	77.0	197679	10	AC128663	AC128663 Mus muscu
C 416	15.4	77.0	114392	2	AC135880	Rattus no	489	15.4	77.0	200500	9	AC093538	AC093538 Homo sapi
C 417	15.4	77.0	116130	9	AC026423	Homo sapi	C 490	15.4	77.0	201106	2	AC110375	AC110375 Mus muscu
C 418	15.4	77.0	120088	9	AC004865	Homo sapi	491	15.4	77.0	201964	10	MMNHG29N7	MMNHG29N7 Mus muscu
C 419	15.4	77.0	126128	10	CNS07YOX	Mus muscu	492	15.4	77.0	201986	2	AC006289	AC006289 Mus muscu
C 420	15.4	77.0	128642	8	AC008838	Homo sapi	C 493	15.4	77.0	202624	10	AC147075	AC147075 Pan trogl
C 421	15.4	77.0	128920	8	AC108871	Oryza sat	494	15.4	77.0	202651	10	AC116587	AC116587 Mus muscu
C 422	15.4	77.0	129355	9	AC004853	Homo sapi	495	15.4	77.0	203363	9	CNS01RHB	AL161668 Human chr
C 423	15.4	77.0	129348	10	CNS07YOZ	Mus muscu	C 496	15.4	77.0	207901	10	AL590389	AL590389 Mouse DNA
C 424	15.4	77.0	134411	8	AL512662	Human DNA	C 497	15.4	77.0	208734	10	AC102675	AC102675 Mus muscu
C 425	15.4	77.0	135378	8	CNS07YPP	Oryza sat	498	15.4	77.0	211554	10	AC099101	AC099101 Rattus no
426	15.4	77.0	136037	8	AC004104	Homo sapi	C 499	15.4	77.0	211931	10	AC125329	AC125329 Mus muscu
C 427	15.4	77.0	136777	2	AC146986	Strongylo	C 500	15.4	77.0	211996	10	AC120837	AC120837 Mus muscu
C 428	15.4	77.0	139266	2	AC146775	Medicago	501	15.4	77.0	213491	2	AC128375	AC128375 Rattus no
429	15.4	77.0	141428	9	AL139415	Human DNA	C 502	15.4	77.0	213634	2	AC123350	AC123350 Mus muscu
430	15.4	77.0	143812	2	AC021847	Homo sapi	503	15.4	77.0	214225	2	AC121430	AC121430 Rattus no

C 504	15.4	77.0	216568	2	AL7J1715	AL7J1715 Mus muscu	C 577	15.2	76.0	1035	6	AX349695	AX349695 Sequence
C 505	15.4	77.0	216957	2	AC119342	AC119342 Rattus no	C 578	15.2	76.0	1273	10	AB008911	AB008911 Mus muscu
C 506	15.4	77.0	219515	2	BX323558	BX323558 Danio rer	C 579	15.2	76.0	1287	6	AX122212	AX122212 Sequence
C 507	15.4	77.0	220807	2	AC139643	AC139643 Rattus no	C 580	15.2	76.0	1287	6	BD164329	BD164329 Novel pol
C 508	15.4	77.0	221952	2	AC098902	AC098902 Rattus no	C 581	15.2	76.0	1332	10	RNPNTS1	RNPNTS1
C 509	15.4	77.0	223102	2	AC094457	AC094457 Rattus no	C 582	15.2	76.0	1335	6	AX067039	AX067039 Sequence
C 510	15.4	77.0	224498	2	AC131973	AC131973 Mus muscu	C 583	15.2	76.0	1410	6	AX067037	AX067037 Sequence
C 511	15.4	77.0	225082	2	AC102806	AC102806 Mus muscu	C 584	15.2	76.0	1417	6	AX813887	AX813887 Sequence
C 512	15.4	77.0	226711	2	AC127440	AC127440 Rattus no	C 585	15.2	76.0	1419	6	AR388495	AR388495 Sequence
C 513	15.4	77.0	227029	2	AC125976	AC125976 Rattus no	C 586	15.2	76.0	1429	5	CHXKPCA	CHXKPCA
C 514	15.4	77.0	228330	10	AC095281	AC095281 Rattus no	C 587	15.2	76.0	1453	8	AF068934	AF068934
C 515	15.4	77.0	229784	10	ALB45171	ALB45171 Mouse DNA	C 588	15.2	76.0	1527	8	AF035700	AF035700
C 516	15.4	77.0	231175	2	AC094467	AC094467 Rattus no	C 589	15.2	76.0	1557	14	AF138298	AF138298
C 517	15.4	77.0	232225	10	AL662876	AL662876 Mouse DNA	C 590	15.2	76.0	1659	6	E03827	E03827
C 518	15.4	77.0	234477	2	AC129063	AC129063 Rattus no	C 591	15.2	76.0	1751	8	AP096247	AP096247
C 519	15.4	77.0	234528	2	AC096222	AC096222 Rattus no	C 592	15.2	76.0	1827	8	BT001108	BT001108
C 520	15.4	77.0	239308	2	AC107111	AC107111 Rattus no	C 593	15.2	76.0	1852	8	AY133701	AY133701
C 521	15.4	77.0	239559	2	AC125908	AC125908 Rattus no	C 594	15.2	76.0	2000	6	AX510247	AX510247
C 522	15.4	77.0	239570	2	AC137983	AC137983 Mus muscu	C 595	15.2	76.0	2000	6	AX652279	AX652279
C 523	15.4	77.0	239726	2	AC097728	AC097728 Rattus no	C 596	15.2	76.0	2161	1	AF411142	AF411142
C 524	15.4	77.0	241587	2	AC096197	AC096197 Rattus no	C 597	15.2	76.0	2180	9	AK126453	AK126453
C 525	15.4	77.0	242686	2	AC103159	AC103159 Rattus no	C 598	15.2	76.0	2317	8	AP170172	AP170172
C 526	15.4	77.0	247978	2	AC094450	AC094450 Rattus no	C 599	15.2	76.0	2319	10	MUSNFKBPA	MUSNFKBPA
C 527	15.4	77.0	248552	2	AC095542	AC095542 Rattus no	C 600	15.2	76.0	2368	1	BD218740	BD218740
C 528	15.4	77.0	253279	2	AC093477	AC093477 Mus muscu	C 601	15.2	76.0	2385	6	BRLCHOB	BRLCHOB
C 529	15.4	77.0	254295	2	BX663526	BX663526 Gallus ga	C 602	15.2	76.0	2385	1	E03828	E03828
C 530	15.4	77.0	254474	2	AC093478	AC093478 Mus muscu	C 603	15.2	76.0	2435	10	S66556	S66556
C 531	15.4	77.0	256779	2	AC105578	AC105578 Rattus no	C 604	15.2	76.0	2483	10	S89033	S89033
C 532	15.4	77.0	257239	2	AC115215	AC115215 Rattus no	C 605	15.2	76.0	2660	14	AY102920	AY102920
C 533	15.4	77.0	257405	2	AC135671	AC135671 Mus muscu	C 606	15.2	76.0	2799	10	RNPEANMT	RNPEANMT
C 534	15.4	77.0	257609	2	AC137477	AC137477 Rattus no	C 607	15.2	76.0	2810	9	BC021128	BC021128
C 535	15.4	77.0	257959	2	AC094598	AC094598 Rattus no	C 608	15.2	76.0	2872	9	BC035072	BC035072
C 536	15.4	77.0	259661	2	AC122570	AC122570 Rattus no	C 609	15.2	76.0	2915	6	AX695502	AX695502
C 537	15.4	77.0	260625	2	AC096018	AC096018 Rattus no	C 610	15.2	76.0	2932	10	AB118195	AB118195
C 538	15.4	77.0	261665	2	AC095565	AC095565 Rattus no	C 611	15.2	76.0	3324	6	AX675131	AX675131
C 539	15.4	77.0	263127	2	AC112337	AC112337 Rattus no	C 612	15.2	76.0	3324	8	HSN802237	HSN802237
C 540	15.4	77.0	263388	2	AC094260	AC094260 Rattus no	C 613	15.2	76.0	3573	8	CHPSRPL2	CHPSRPL2
C 541	15.4	77.0	271135	2	AC103217	AC103217 Mus muscu	C 614	15.2	76.0	3657	6	AX675130	AX675130
C 542	15.4	77.0	274751	2	BX649574	BX649574 Mus muscu	C 615	15.2	76.0	3663	6	AX675129	AX675129
C 543	15.4	77.0	275392	2	AC095297	AC095297 Rattus no	C 616	15.2	76.0	3762	6	AX675147	AX675147
C 544	15.4	77.0	278425	2	AC103101	AC103101 Rattus no	C 617	15.2	76.0	3768	6	AX675146	AX675146
C 545	15.4	77.0	288495	2	AC095608	AC095608 Rattus no	C 618	15.2	76.0	3892	6	AX306113	AX306113
C 546	15.4	77.0	291709	2	AC094701	AC094701 Rattus no	C 619	15.2	76.0	3892	6	AX695501	AX695501
C 547	15.4	77.0	294316	2	AC131613	AC131613 Rattus no	C 620	15.2	76.0	3892	10	MUSPEBDS	MUSPEBDS
C 548	15.4	77.0	299859	2	AC128993	AC128993 Rattus no	C 621	15.2	76.0	3912	1	REO242746	REO242746
C 549	15.4	77.0	300892	1	AE016747	AE016747 Staphyloc	C 622	15.2	76.0	4114	10	AF110477	AF110477 Rattus no
C 550	15.4	77.0	302835	1	AE012555	AE012555 Xylella f	C 623	15.2	76.0	4270	3	AF070482	AF070482
C 551	15.4	77.0	311857	2	BX000982	BX000982 Mus muscu	C 624	15.2	76.0	4304	10	AF110478	AF110478
C 552	15.4	77.0	319072	3	AE003474	AE003474 Drosophila	C 625	15.2	76.0	4330	9	AB052103	AB052103
C 553	15.4	77.0	335913	6	AX196295	AX196295 Sequence	C 626	15.2	76.0	4591	14	ASU13763	ASU13763
C 554	15.4	77.0	335913	6	AX196295	AX196295 Sequence	C 627	15.2	76.0	4884	6	AR319928	AR319928
C 555	15.2	76.0	170	14	VSVDLRA	M10402 VSV defecti	C 628	15.2	76.0	5321	8	AF116598	AF116598
C 556	15.2	76.0	177	14	RHVSU06	XO0464 5'-region o	C 629	15.2	76.0	5327	6	AX100077	AX100077
C 557	15.2	76.0	259	11	G63946	G63946 UMBTL003 Bo	C 630	15.2	76.0	5455	2	AX069012	AX069012
C 558	15.2	76.0	259	11	KLAJ9990	ALJ229990 Kluverrom	C 631	15.2	76.0	5661	9	AF113514	AF113514
C 559	15.2	76.0	308	6	AX920060	AX920060 Sequence	C 632	15.2	76.0	5881	6	AX807249	AX807249
C 560	15.2	76.0	308	6	BD055593	BD055593 Sequence	C 633	15.2	76.0	5881	6	AX807415	AX807415
C 561	15.2	76.0	330	6	AR355019	AR355019 Sequence	C 634	15.2	76.0	5909	9	AB061870	AB061870
C 562	15.2	76.0	366	14	VSVDGL	K03386 Vesticular s	C 635	15.2	76.0	5988	9	AF119230	AF119230
C 563	15.2	76.0	422	11	G24310	G24310 human STS W	C 636	15.2	76.0	6222	6	AX695649	AX695649
C 564	15.2	76.0	441	14	VSVDLITG	VSVDLITG	C 637	15.2	76.0	6330	6	AX100076	AX100076
C 565	15.2	76.0	443	6	AX675141	AX675141 Sequence	C 638	15.2	76.0	6380	6	AX100079	AX100079
C 566	15.2	76.0	450	8	AX587452	AX587452 Arabidops	C 639	15.2	76.0	6380	14	RHVSU06	RHVSU06
C 567	15.2	76.0	499	6	AX675139	AX675139 Sequence	C 640	15.2	76.0	6537	6	AX695648	AX695648
C 568	15.2	76.0	548	6	AX675140	AX675140 Sequence	C 641	15.2	76.0	6537	9	AF119231	AF119231
C 569	15.2	76.0	581	6	AX675137	AX675137 Sequence	C 642	15.2	76.0	6592	9	AB002381	AB002381
C 570	15.2	76.0	578	11	AX675136	AX675136 Sequence	C 643	15.2	76.0	8293	9	AF217500	AF217500
C 571	15.2	76.0	581	11	BV017953	BV017953 S209P6355	C 644	15.2	76.0	8293	9	DMU77947	DMU77947
C 572	15.2	76.0	608	11	BV043721	BV043721 S212P6535	C 645	15.2	76.0	8401	3	AY095952	AY095952
C 573	15.2	76.0	632	11	BV070174	BV070174 S212P6118	C 646	15.2	76.0	8877	1	AY095952	AY095952
C 574	15.2	76.0	677	14	VSVDLITL	K01640 Vesticular s	C 647	15.2	76.0	9196	6	G07GRB2E5	G07GRB2E5
C 575	15.2	76.0	708	6	AX675135	AX675135 Sequence	C 648	15.2	76.0	10618	2	AC015408	AC015408
C 576	15.2	76.0	859	11	BV065920	BV065920 S212P6911	C 649	15.2	76.0	11161	14	VSVCG	VSVCG

650	15.2	76.0	11162	14	AF473864	AF473864 Vesticular	C 723	15.2	76.0	11698	10	AL731687	AL731687 Mouse DNA
651	15.2	76.0	11167	1	AE001599	AE001599 Chlamydia	C 724	15.2	76.0	116951	8	AC114788	AC114788 Homo sapi
652	15.2	76.0	11168	1	AE002222	AE002222 Chlamydia	C 725	15.2	76.0	117226	9	AP006626	AP006626 Oryza sat
653	15.2	76.0	11688	1	AE009887	AE009887 Pyrobacul	C 726	15.2	76.0	118235	10	AC003060	AC003060 Mus muscu
654	15.2	76.0	12132	1	AE008393	AE008393 Streptoco	C 727	15.2	76.0	118313	9	AC003991	AC003991 Human BAC
655	15.2	76.0	12132	8	AB013613	AB013613 Oryza sat	C 728	15.2	76.0	119666	9	HSDJ1B2	HSDJ1B2 Homo sapi
656	15.2	76.0	14179	1	AE007053	AE007053 Mycobacte	C 729	15.2	76.0	121961	8	AC090556	AC090556 Homo sapi
657	15.2	76.0	14311	6	AR123015	AR123015 Sequence	C 730	15.2	76.0	122151	10	NCB23H20	NCB23H20 Neurospor
658	15.2	76.0	14311	6	AR123016	AR123016 Sequence	C 731	15.2	76.0	122356	10	AL845310	AL845310 Mouse DNA
659	15.2	76.0	16332	2	AC023615	AC023615 Mus muscu	C 732	15.2	76.0	123039	9	AC093087	AC093087 Homo sapi
660	15.2	76.0	19300	14	AF327841	AF327841 African s	C 733	15.2	76.0	123954	8	AC123518	AC123518 Oryza sat
661	15.2	76.0	20868	2	AC017847	AC017847 Drosophill	C 734	15.2	76.0	125008	10	AB005546	AB005546 Homo sapi
662	15.2	76.0	23009	1	AE008763	AE008763 Salmonell	C 735	15.2	76.0	127220	10	BX284649	BX284649 Mouse DNA
663	15.2	76.0	25967	2	AC100971	AC100971 Mus muscu	C 736	15.2	76.0	127334	9	AC109579	AC109579 Homo sapi
664	15.2	76.0	30875	2	HSN20A6	Z69173 Human DNA s	C 737	15.2	76.0	127790	9	AC008885	AC008885 Homo sapi
665	15.2	76.0	34051	9	AC104668	AC104668 Homo sapi	C 738	15.2	76.0	128580	2	AP005466	AP005466 Oryza sat
666	15.2	76.0	34713	2	BX539329	BX539329 Homo sapi	C 739	15.2	76.0	128683	9	AC027329	AC027329 Homo sapi
667	15.2	76.0	36411	1	U000022	U000022 Mycobacteri	C 740	15.2	76.0	128757	9	AL359837	AL359837 Human DNA
668	15.2	76.0	36412	6	AR345364	AR345364 Sequence	C 741	15.2	76.0	128856	2	AC126779	AC126779 Medicago
669	15.2	76.0	38322	14	AF327839	AF327839 African s	C 742	15.2	76.0	131098	2	AC146865	AC146865 Medicago
670	15.2	76.0	39089	3	CBRG33D04	AC084550 Caenorhab	C 743	15.2	76.0	131427	14	OPU75930	OPU75930 Homo sapi
671	15.2	76.0	40520	9	AC011513	AC011513 Homo sapi	C 744	15.2	76.0	131995	9	AC008496	AC008496 Homo sapi
672	15.2	76.0	40866	9	AL590393	AL590393 Human DNA	C 745	15.2	76.0	135434	9	AL358474	AL358474 Homo sapi
673	15.2	76.0	43593	9	AC022516	AC022516 Homo sapi	C 746	15.2	76.0	137347	9	AL358674	AL358674 Homo sapi
674	15.2	76.0	44806	9	AC069162	AC069162 Homo sapi	C 747	15.2	76.0	137908	10	AL607024	AL607024 Homo sapi
675	15.2	76.0	45289	3	CBRG39N24	AC084576 Caenorhab	C 748	15.2	76.0	138070	2	AP002028	AP002028 Homo sapi
676	15.2	76.0	47338	10	AF326469	AF326469 Mus muscu	C 749	15.2	76.0	138621	2	AC104550	AC104550 Mus muscu
677	15.2	76.0	47506	2	AC018082	AC018082 Drosophill	C 750	15.2	76.0	138982	2	AC066602	AC066602 Homo sapi
678	15.2	76.0	49336	1	AC025950	AC025950 Staphyloc	C 751	15.2	76.0	139443	10	AC124376	AC124376 Mus muscu
679	15.2	76.0	50152	1	AC025948	AC025948 Staphyloc	C 752	15.2	76.0	139636	2	AP005558	AP005558 Oryza sat
680	15.2	76.0	51410	9	AC092604	AC092604 Homo sapi	C 753	15.2	76.0	141428	2	AC121253	AC121253 Homo sapi
681	15.2	76.0	53886	2	AC145962	AC145962 Gallus ga	C 754	15.2	76.0	142193	2	AC025364	AC025364 Homo sapi
682	15.2	76.0	54157	8	T7A14	AC005532 Arabidops	C 755	15.2	76.0	142291	2	AC079578	AC079578 Mus muscu
683	15.2	76.0	60801	9	AC006551	AC006551 Homo sapi	C 756	15.2	76.0	143384	2	AC129050	AC129050 Rattus no
684	15.2	76.0	61579	9	AL356258	AL356258 Human DNA	C 757	15.2	76.0	143570	8	EX005168	EX005168 Mouse DNA
685	15.2	76.0	62794	8	AP003957	AP003957 Oryza sat	C 758	15.2	76.0	143588	10	AC106793	AC106793 Homo sapi
686	15.2	76.0	63501	2	AC134974	AC134974 Homo sapi	C 759	15.2	76.0	146040	10	AC012399	AC012399 Homo sapi
687	15.2	76.0	63604	8	AP000381	AP000381 Arabidops	C 760	15.2	76.0	148383	2	CNS01DSL	AC112819 Human chr
688	15.2	76.0	69019	2	AC005117	AC005117 Drosophill	C 761	15.2	76.0	148483	2	AC091701	AC091701 Trypanoso
689	15.2	76.0	70872	2	AC114997	AC114997 Mus muscu	C 762	15.2	76.0	149483	2	AC100744	AC100744 Mus muscu
690	15.2	76.0	71044	2	AC024368	AC024368 Homo sapi	C 763	15.2	76.0	149651	8	AF229187	AF229187 Oryza sat
691	15.2	76.0	74512	2	AC100521	AC100521 Mus muscu	C 764	15.2	76.0	149746	2	AC116245	AC116245 Rattus no
692	15.2	76.0	74675	8	AP004787	AP004787 Oryza sat	C 765	15.2	76.0	149854	10	AC012399	AC012399 Homo sapi
693	15.2	76.0	78507	2	AC019704	AC019704 Drosophill	C 766	15.2	76.0	149871	2	BX842689	AC112210 Homo sapi
694	15.2	76.0	79877	2	AC131495	AC131495 Lytechinu	C 767	15.2	76.0	146699	2	AC116874	AC116874 Homo sapi
695	15.2	76.0	82059	9	AC130896	AC130896 Homo sapi	C 768	15.2	76.0	147746	2	AC116874	AC116874 Homo sapi
696	15.2	76.0	82080	10	AL844553	AL844553 Mouse DNA	C 769	15.2	76.0	147946	2	AC091432	AC091432 Homo sapi
697	15.2	76.0	82933	6	AX704274	AX704274 Sequence	C 770	15.2	76.0	147946	2	AL772207	AL772207 Mouse DNA
698	15.2	76.0	85339	9	AL139086	AL139086 Human DNA	C 771	15.2	76.0	148320	10	AL772207	AL772207 Mouse DNA
699	15.2	76.0	85641	9	AC005217	AC005217 Homo sapi	C 772	15.2	76.0	148325	5	EX005124	EX005124 Zebrafish
700	15.2	76.0	86146	2	AC0023805	AC0023805 Mus muscu	C 773	15.2	76.0	149748	2	AC138287	AC138287 Mus muscu
701	15.2	76.0	88401	8	F23M19	AC007455 Arabidops	C 774	15.2	76.0	149844	2	AC144682	AC144682 Rattus no
702	15.2	76.0	91825	2	AC019854	AC019854 Drosophill	C 775	15.2	76.0	149995	9	AC016722	AC016722 Homo sapi
703	15.2	76.0	92562	10	AL732485	AL732485 Mouse DNA	C 776	15.2	76.0	150292	10	AC112342	AC112342 Rattus no
704	15.2	76.0	96079	9	HS717L17	AL139086 Human DNA	C 777	15.2	76.0	150855	2	AC021258	AC021258 Homo sapi
705	15.2	76.0	96598	6	AX695500	AX695500 Sequence	C 778	15.2	76.0	151073	2	AC134624	AC134624 Oryza sat
706	15.2	76.0	97902	9	AL596118	AL596118 Human DNA	C 779	15.2	76.0	151289	2	AC102184	AC102184 Mus muscu
707	15.2	76.0	106123	9	AL158015	AL158015 Human DNA	C 780	15.2	76.0	151559	9	AL390294	AL390294 Human DNA
708	15.2	76.0	109770	2	AC073763_3	Continuaction (4 of	C 781	15.2	76.0	152152	2	AC113505	AC113505 Mus muscu
709	15.2	76.0	110000	2	AC095248_0	AC095248 Rattus no	C 782	15.2	76.0	152304	2	AC023032	AC023032 Homo sapi
710	15.2	76.0	110000	2	AC096079_2	Continuaction (3 of	C 783	15.2	76.0	152823	2	AC010210	AC010210 Homo sapi
711	15.2	76.0	110000	2	AC098642_2	Continuaction (3 of	C 784	15.2	76.0	153080	2	AC026229	AC026229 Homo sapi
712	15.2	76.0	110000	2	AC110355_2	Continuaction (3 of	C 785	15.2	76.0	153520	10	AC122520	AC122520 Mus muscu
713	15.2	76.0	110000	2	AC116130_1	Continuaction (2 of	C 786	15.2	76.0	153994	8	AC138023	AC138023 Homo sapi
714	15.2	76.0	110000	2	BM572642_03	Continuaction (3 of	C 787	15.2	76.0	155634	9	AP000391	AP000391 Oryza sat
715	15.2	76.0	110000	2	LMFLCHR25_03	Continuaction (4 of	C 788	15.2	76.0	155785	10	AC117262	AC117262 Mus muscu
716	15.2	76.0	110000	6	AR310754_01	Continuaction (2 of	C 789	15.2	76.0	157394	2	AP004874	AP004874 Oryza sat
717	15.2	76.0	110102	8	ATP28J12	AL021710 Arabidops	C 790	15.2	76.0	157994	9	AC114478	AC114478 Homo sapi
718	15.2	76.0	111354	8	AP003806	AP003806 Oryza sat	C 791	15.2	76.0	158276	9	AL160267	AL160267 Human DNA
719	15.2	76.0	111568	9	AP003908	AP003908 Homo sapi	C 792	15.2	76.0	158315	9	AC114297	AC114297 Homo sapi
720	15.2	76.0	112855	2	AP003608	AP003608 Oryza sat	C 793	15.2	76.0	158406	9	AC114297	AC114297 Homo sapi
721	15.2	76.0	115289	9	AC011490	AC011490 Homo sapi	C 794	15.2	76.0	159070	2	AC128114	AC128114 Rattus no
722	15.2	76.0	116653	2	AP005574	AP005574 Oryza sat	C 795	15.2	76.0	159104	9	AC108925	AC108925 Homo sapi

796 15.2 76.0 159468 2 AC023314 Homo sapi
c 797 15.2 76.0 159519 2 AC018511 Homo sapi
c 798 15.2 76.0 159617 9 AC022018 Homo sapi
799 15.2 76.0 159832 2 AC104199 Mus muscu
c 800 15.2 76.0 160039 2 AC104851 Homo sapi

ALIGNMENTS

RESULT 1
AX282880 20 bp DNA linear PAT 02-NOV-2001
LOCUS Sequence 3 from Patent WO0174346.
AX282880
ACCESSION AX282880.1 GI:16609856
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

1
Agrawal, S. and Kandimala, E.R.
Sensitization of cells to cytotoxic agents using oligonucleotides
directed to nucleotide excision repair or transcription coupled
repair genes
Patent: WO 0174346-A 3 11-OCT-2001;
HYBRIDON, INC. (US)
location/Qualifiers

FEATURES
source
1..20
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="HYB 963 oligonucleotide"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTCCATCTCATGTTGATG 20
Db 1 GGTCCATCTCATGTTGATG 20

RESULT 2
HSXPAC5 809 bp DNA linear PRI 27-MAR-1996
LOCUS Human Xeroderma Pigmentosum group A correcting (XPAC) gene, exon 6
DEFINITION
ACCESSION
VERSION
KEYWORDS
SEGMENT
SOURCE
ORGANISM

5 of 5
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 809)
Topping, R.S., Myrand, S.P., Williams, B.L., Albert, J.C. and
States, J.C.
Characterization of the human XPA promoter
Gene 166 (2), 341-342 (1995)
MEDLINE 96125219
PUBMED 8543191

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
1..809
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/clone="Lambda XPAC 1"
/sex="male"
/tissue_type="placental"
/clone_lib="Lambda FIX II Genomic DNA library from
Stratagene"
Join(U10343.1:88..298,U10344.1:1..306,U10345.1:1..366,
U10346.1:1..318,1..754)
/gene="XPAC"
Join(U10343.1:<88..198,U10344.1:101..206,
U10345.1:101..266,U10346.1:101..218,77..225)
/gene="XPAC"

CDS

exon

polyA_signal
polyA_signal
polyA_signal
/gene="XPAC"
/number=6
267..272
/gene="XPAC"
729..734
/gene="XPAC"

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 809;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTCCATCTCATGTTGATG 20
Db 146 GGTCCATCTCATGTTGATG 127

RESULT 3
AR183106/c 822 bp DNA linear PAT 20-APR-2002
LOCUS AR183106
DEFINITION Sequence 8 from patent US 6340566.
ACCESSION AR183106
VERSION AR183106.1 GI:20226699
KEYWORDS
SOURCE
ORGANISM

Unknown.
Unclassified.
1 (bases 1 to 822)
McCutchen-Malone, S.L.
Detection and quantitation of single nucleotide polymorphisms, DNA
sequence variations, DNA mutations, DNA damage and DNA mismatches
Patent: US 6340566-A 8 22-JAN-2002;
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
1..822
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 822;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTCCATCTCATGTTGATG 20
Db 743 GGTCCATCTCATGTTGATG 724

RESULT 4
AR203243/c 822 bp DNA linear PAT 20-JUN-2002
LOCUS AR203243
DEFINITION Sequence 8 from patent US 6365355.
ACCESSION AR203243

VERSION	AR203243.1	GI:21499581
KEYWORDS	.	Unknown.
SOURCE	.	Unknown.
ORGANISM	.	Unclassified.
REFERENCE	1	(bases 1 to 822)
AUTHORS	Mccutchen-Maloney,S.L.	
TITLE	Chimeric proteins for detection and quantitation of DNA mutations, DNA sequence variations, DNA damage and DNA mismatches	
JOURNAL	Patent: US 636535-A 8 02-APR-2002;	
FEATURES	location/Qualifiers	
source	1..822	
	/organism="unknown"	
	/mol_type="unassigned DNA"	
ORIGIN		
Query Match	100.0%; Score 20; DB 6; Length 822;	
Best Local Similarity	100.0%; Pred. No. 15;	
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 GGTCATCTCATGTTGATG 20	
Db	743 GGTCATCTCATGTTGATG 724	
RESULT 5		
LOCUS	AX277000 822 bp DNA linear PAT 29-OCT-2001	
DEFINITION	Sequence 8 from Patent WO0173079.	
ACCESSION	AX277000	
VERSION	AX277000.1 GI:16548670	
KEYWORDS	.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1	
AUTHORS	Mccutchen-Maloney,S.L.	
TITLE	Chimeric proteins for detection and quantitation of dna mutations, dna sequence variations, dna damage and dna mismatches	
JOURNAL	Patent: WO 0173079-A 8 04-OCT-2001;	
FEATURES	The Regents of The University of California (US)	
source	location/Qualifiers	
	1..822	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
Query Match	100.0%; Score 20; DB 6; Length 822;	
Best Local Similarity	100.0%; Pred. No. 15;	
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 GGTCATCTCATGTTGATG 20	
Db	743 GGTCATCTCATGTTGATG 724	
RESULT 6		
LOCUS	AX282890 1377 bp DNA linear PAT 02-NOV-2001	
DEFINITION	Sequence 13 from Patent W00174346.	
ACCESSION	AX282890	
VERSION	AX282890.1 GI:16609866	
KEYWORDS	.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1	
AUTHORS	Agrawal,S. and Kandimala,E.R.	
TITLE	Sensitization of cells to cytotoxic agents using oligonucleotides directed to nucleotide excision repair or transcription coupled	

JOURNAL	repair genes
PATENT	Patent: WO 0174346-A 13 11-OCT-2001;
HYBRIDION	HYBRIDION, INC. (US)
FEATURES	Location/Qualifiers
SOURCE	1..1377
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	100.0%; Score 20; DB 6; Length 1377;
Best Local Similarity	100.0%; Pred. No. 16;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GGTCATCTCATCTGTGATG 20
Db	769 GGTCATCTCATCTGTGATG 750
RESULT 7	
LOCUS	HUMXPAC 1377 bp mRNA linear PRI 11-FEB-2003
DEFINITION	Homo sapiens mRNA for XPAC protein, complete cds.
ACCESSION	D14533
VERSION	D14533.1 GI:286028
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 1377)
AUTHORS	Tanaka, K., Miura, N., Satokata, I., Miyamoto, I., Yoshida, M.C.,
	Sato, Y., Kondo, S., Yasui, A., Okayama, H. and Okada, Y.
	Analysis of a human DNA excision repair gene involved in group A
	xeroderma pigmentosum and containing a zinc-finger domain
	Nature 348 (6296), 73-76 (1990)
TITLE	91043046
JOURNAL	2234061
MEDLINE	
PUBMED	
REFERENCE	2
AUTHORS	Satokata, I., Iwai, K., Matsuda, T., Okada, Y. and Tanaka, K.
TITLE	Genomic characterization of the human DNA excision
	repair-controlling gene XPAC
	Gene 136 (1-2), 345-348 (1993)
JOURNAL	94124028
MEDLINE	
PUBMED	8294029
REFERENCE	3 (bases 1 to 1377)
AUTHORS	Tanaka, K.
TITLE	Direct Submission
JOURNAL	Submitted (26-FEB-1993) Kiyoi Tanaka, Osaka University, Inst. for
	Molecular and Cellular Biology; 1-3 Yamadaoka, Suita, Osaka 565,
	Japan (Tel:81-6-6877-5238, Fax:81-6-6877-9136)
FEATURES	Location/Qualifiers
SOURCE	1..1377
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="9"
	/clone_id="pcd2Basinger"
	/note="clones pcd2h19 and pcd2h29-human primer y
	fibroblast"
	1..1377
gene	/gene="XPAC"
	27..848
CDS	/gene="XPAC"
	/codon_start=1
	/product="XPAC protein"
	/protein_id="BAA03403.1"
	/db_xref="GI:286029"
	/translation="MAADAGLPEPAALLEGPAELPASVRASIERKQPALMLROARLA
	ARPSATAAATGMAVVAAPKIDTGGGFIIEEBEERQIKGVHQPVAFVY
	VICEGCEPMDSVLMNHFDLPTCDNRDADDKHLITTEAKOEYLLKDDLEKREBE
	PLKATVKNPHSOWGMKYLKQIYKRSLEWGSQEALEAEKEVQENRKKOKK
	FDKVKELRAVRSSVMKRETIYHQHEYGPENLDDMYRTCTMCGHELTYEKM"

repeat_region /rpt_type=dispersed
210. .431
/rpt_family="Alu"
/rpt_type=dispersed
variation /frequency="0.13"
/replace="t"
395
/frequency="0.15"
/replace="g"
repeat_region 433. .493
/rpt_family="L1"
/rpt_type=dispersed
variation 535
/frequency="0.01"
/replace="a"
600
/frequency="0.16"
/replace="g"
variation 701
/frequency="0.01"
/replace="c"
706
/frequency="0.15"
/replace="t"
repeat_region 1013. .1076
/rpt_family="MIR"
/rpt_type=dispersed
variation 1168
/frequency="0.14"
/replace="t"
1169. .1381
/rpt_family="MIR"
/rpt_type=dispersed
variation 1428
/frequency="0.01"
/replace="t"
1536. .1853
/rpt_family="L1"
/rpt_type=dispersed
variation 1595
/frequency="0.01"
/replace="c"
<1643. .24097
/gene="XPA"
join(<1643. .1840,5207. .5317,9329. .9434,11710. .11875,
13931. .14048,23420. .24097)
/gene="XPA"
/product="xeroderma pigmentosum, complementation group A"
1665
/gene="XPA"
/frequency="0.38"
/replace="a"
join(1669. .1840,5207. .5317,9329. .9434,11710. .11875,
13931. .14048,23420. .23568)
/gene="XPA"
/codon_start=1
/product="xeroderma pigmentosum, complementation group A"
/protein_id="AA118969.1"
/cd_xref="GI:20303098"
/translation="MAADGALPEAALDEQPAELPASVASIERKQRLMRLQRLA
ARPYSAATAATGMAVAAPKIIDTGGFTLEEEBEEKIGKVHQPQGVMEFDY
VICCEGKEFMSYLMNHFDLPDCNRDADDKKLITTEAKQETLKDCDLKEEP
PLKEITVKRNPHSOMGDKLYLKLOI VKRSLEWGSQALTEAKKEVROENREKQKK
FDKKVLELRRAVSSVWKRETTVHQHEFYEPENLEDDMRKTCMGHELTTEKM"
1934
/gene="XPA"
/frequency="0.15"
/replace="c"
2044
/gene="XPA"
/frequency="0.13"
/replace="g"

repeat_region 2182. .2486
/rpt_family="CR1"
/rpt_type=dispersed
variation 2195
/gene="XPA"
/frequency="0.08"
/replace="a"
2424
/gene="XPA"
/frequency="0.02"
/replace="g"
variation 2680
/gene="XPA"
/frequency="0.15"
/replace="a"
2903
/gene="XPA"
/frequency="0.16"
/replace="g"
variation 3134
/gene="XPA"
/frequency="0.15"
/replace="t"
3173
/gene="XPA"
/frequency="0.02"
/replace="g"
3484
/gene="XPA"
/frequency="0.03"
/replace="c"
3509
/gene="XPA"
/frequency="0.05"
/replace="g"
3649. .3706
/rpt_family="MIR"
/rpt_type=dispersed
variation 3738. .3742
/gene="XPA"
/frequency="0.85"
/replace="t"
3785. .4598
/rpt_family="L1"
/rpt_type=dispersed
variation 3818
/gene="XPA"
/frequency="0.11"
/replace="a"
4188
/gene="XPA"
/frequency="0.05"
/replace="a"
4487
/gene="XPA"
/frequency="0.04"
/replace="t"
4607. .4694
/rpt_family="MER2_type"
/rpt_type=dispersed
variation 4651
/gene="XPA"
/frequency="0.02"
/replace="t"
4710
/gene="XPA"
/frequency="0.02"
/replace="t"
4711. .4714
/gene="XPA"
/frequency="0.06"
/replace="t"
repeat_region 4738. .5108

```

variation      /rpt_family="L1"
                /rpt_type="dispersed"
                /gene="XPA"
                /frequency="0.02"
                /replace="c"
variation      4839
                /gene="XPA"
                /frequency="0.16"
                /replace="t"
variation      5272..5274
                /gene="XPA"
                /note="deletion results in removal of glutamate residue
                /from amino acid sequence"
                /frequency="0.01"
                /replace="n"
variation      5522
                /gene="XPA"
                /frequency="0.01"
                /replace="a"
repeat_region  /rpt_family="CR1"
                /rpt_type="dispersed"
                /gene="XPA"
                /note="Region not scanned for variation"
                /frequency="0.001"
                /replace="n"

Query Match    100.0%; Score 20; DB 9; Length 24994;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTACTCAGTGTGATG 20
    |||||
Db 23489 GGCCTACTCAGTGTGATG 23470

RESULT 10
AL445531 111345 bp DNA linear PRI 01-MAR-2001
LOCUS      Human DNA sequence from clone RP11-54606 on chromosome 9, complete
            sequence.
ACCESSION  AL445531
VERSION    AL445531.10 GI:13234986
KEYWORDS   HTG.
SOURCE      Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 111345)
REFERENCE   Hammond S.
            Direct Submissiion
            Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
            On Mar 5, 2001 this sequence version replaced gi:11184378.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em: EMBL; Sw:
            SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
            database can be found at

```

```

FEATURES
    source
        http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
        was generated from part of bacterial clone contigs of human
        chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
        Group. Further information can be found at
        http://www.sanger.ac.uk/HGP/Chr9
        RP11-54606 is from the library RPCT-11.2 constructed by the group
        of Pieter de Jong. For further details see
        http://www.chori.org/bacpac/home.htm
        VECTOR: pBAC63.6
        IMPORTANT: This sequence is not the entire insert of clone
        RP11-54606. It may be shorter because we sequence overlapping
        sections only once, except for a 100 base overlap.
        The true left end of clone RP11-23B15 is at 111246 in this
        sequence. The true right end of clone RP11-24A49 is at 100 in this
        sequence.

        1..111345
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="9"
        /clone="RP11-54606"
        /clone_11b="RPCT-11.2"
        /note="MER5B repeat: matches 2..151 of consensus"
        2143..2375
        /note="Tigger2a repeat: matches 1..212 of consensus"
        2376..2671
        /note="AluY repeat: matches 1..236 of consensus"
        2672..2836
        /note="Tigger2a repeat: matches 212..434 of consensus"
        3484..3661
        /note="MIR repeat: matches 28..250 of consensus"
        4208..4283
        /note="MIR repeat: matches 181..254 of consensus"
        6161..6299
        /note="MIR repeat: matches 79..233 of consensus"
        6933..7250
        /note="AluJo repeat: matches 1..310 of consensus"
        7601..7647
        /note="T1MC4 repeat: matches 6807..6852 of consensus"
        7648..7954
        /note="AluX repeat: matches 1..307 of consensus"
        7955..8098
        /note="T1MC4 repeat: matches 6852..6999 of consensus"
        8266..8937
        /note="T1MC4 repeat: matches 7297..7976 of consensus"
        8942..9568
        /note="T1MC4 repeat: matches 6854..7538 of consensus"
        9621..9741
        /note="T1MC4 repeat: matches 7625..7757 of consensus"
        9742..10054
        /note="AluSq repeat: matches 1..296 of consensus"
        10055..10134
        /note="T1MC4 repeat: matches 7757..7833 of consensus"
        13726..13782
        /note="U2 repeat: matches 2654..2710 of consensus"
        13932..14026
        /note="MIR repeat: matches 37..141 of consensus"
        16174..16458
        /note="AluDb repeat: matches 1..306 of consensus"
        16989..17178
        /note="U2 repeat: matches 2223..2419 of consensus"
        17214..17390
        /note="U2 repeat: matches 1949..2126 of consensus"
        17428..17503
        /note="U2 repeat: matches 1797..1869 of consensus"
        17706..18041
        /note="MT1D repeat: matches 57..382 of consensus"
        19266..19365
        /note="50 copies 2 mer aa 62% conserved"
        19416..19742
        /note="MT1D repeat: matches 204..505 of consensus"

```

repeat_region	19743..19906	/note="AluSx repeat: matches 139..297 of consensus"
repeat_region	19907..20214	/note="AluSx repeat: matches 1..308 of consensus"
repeat_region	20215..20351	/note="AluSx repeat: matches 1..139 of consensus"
repeat_region	20352..20548	/note="MLTID repeat: matches 1..204 of consensus"
repeat_region	20757..20833	/note="U2 repeat: matches 2576..2661 of consensus"
repeat_region	20954..21262	/note="AluSg repeat: matches 6..303 of consensus"
repeat_region	21263..21416	/note="AluSc repeat: matches 132..285 of consensus"
repeat_region	21417..21456	/note="20 copies 2 mer ta 92% conserved"
repeat_region	21540..21839	/note="AluSp repeat: matches 1..302 of consensus"
repeat_region	22282..22443	/note="Charlie3 repeat: matches 2531..2685 of consensus"
repeat_region	23387..23579	/note="AluI4 repeat: matches 4077..4280 of consensus"
repeat_region	23572..23757	/note="U1M4 repeat: matches 4416..4608 of consensus"
repeat_region	23801..23887	/note="MER44 repeat: matches 637..724 of consensus"
repeat_region	23897..24410	/note="U1M3 repeat: matches 5239..6095 of consensus"
misc_feature	26344..27022	/note="Opc island"
repeat_region	27109..27321	/evidence=not_experimental
repeat_region	27997..28057	/note="MIR repeat: matches 9..250 of consensus"
repeat_region	28059..28280	/note="U1M5 repeat: matches 5098..5164 of consensus"
repeat_region	28313..28395	/note="AluIo repeat: matches 1..220 of consensus"
repeat_region	30327..30625	/note="AluIv repeat: matches 209..291 of consensus"
repeat_region	30684..31515	/note="U1M5 repeat: matches 5152..5456 of consensus"
repeat_region	31516..32221	/note="HERV1_48 repeat: matches 1103..1927 of consensus"
repeat_region	32463..32508	/note="U1M5 repeat: matches 5463..6176 of consensus"
repeat_region	32518..32784	/note="MER5A repeat: matches 61..106 of consensus"
repeat_region	32877..33055	/note="U1M1 repeat: matches 6056..6325 of consensus"
repeat_region	34484..34613	/note="MER5B repeat: matches 5..174 of consensus"
repeat_region	34819..35010	/note="U2 repeat: matches 2576..2710 of consensus"
repeat_region	35519..35707	/note="MIR repeat: matches 5..193 of consensus"
repeat_region	35920..36029	/note="U2 repeat: matches 2006..2204 of consensus"
repeat_region	36429..37325	/note="U2 repeat: matches 2627..2750 of consensus"
repeat_region	37597..37958	/note="U1M8 repeat: matches 5258..6171 of consensus"
repeat_region	38790..38968	/note="HER1B repeat: matches 1..364 of consensus"
repeat_region	39203..39497	/note="AluIo repeat: matches 1..175 of consensus"
repeat_region	41276..41626	/note="AluSc repeat: matches 1..236 of consensus"
repeat_region	41741..42056	/note="U1M3 repeat: matches 5800..6155 of consensus"
repeat_region	42238..42458	/note="AluIo repeat: matches 1..310 of consensus"
repeat_region	42238..42458	/note="MER30 repeat: matches 1..230 of consensus"

Query Match	Best Local Similarity	100.0%;	Score 20;	DB 9;	Length 111345;
Matches	20;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
repeat_region	42810..43302	/note="LIM3 repeat: matches 5555..6074 of consensus"			
repeat_region	43417..43581	/note="Charli4 repeat: matches 238..1961 of consensus"			
repeat_region	43894..44020	/note="WIR repeat: matches 62..188 of consensus"			
repeat_region	44466..44760	/note="ALusg repeat: matches 2..306 of consensus"			
repeat_region	45611..45897	/note="ALusg repeat: matches 1..307 of consensus"			
repeat_region	46001..46306	/note="ALusg repeat: matches 1..312 of consensus"			
repeat_region	47234..47340	/note="ALusg repeat: matches 1..301 of consensus"			
repeat_region	48040..48297	/note="WIR repeat: matches 28..139 of consensus"			
repeat_region	48578..48710	/note="ALusg repeat: matches 38..295 of consensus"			
repeat_region	48710..48810	/note="FLAM_C repeat: matches 1..133 of consensus"			
Query Match	100.0%;	Score 20;	DB 9;	Length 111345;	
Best Local Similarity	100.0%;	Pred. No. 20;			
Matches	20;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	GGTCATACATCATGTTGATG	20		
Db	5047	GGTCATACATCATGTTGATG	5066		
RESULT 11					
LOCUS	AL442130/c	187079 bp	DNA	linear	HTG 10-JUL-2001
DEFINITION	Homo sapiens chromosome 9 clone RP11-24E19, 7 unordered pieces.				
ACCESSION	AL442130				
VERSION	AL442130.2	GI:10716422			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_CANCELLED.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1	Plumb, B.			
AUTHORS	Direct Submission				
TITLE	Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk				
JOURNAL	Requests: clonerequest@sanger.ac.uk				
	On Oct 7, 2000 this sequence version replaced gi:10278028.				
COMMENT	Genome Center				
	Center: Sanger Centre				
	Center code: SC				
	Web site: http://www.sanger.ac.uk				
	Contact: humquery@sanger.ac.uk				
	Project Information				
	Center project name: ba24E19				
	Summary Statistics				
	Assembly program: XGAP4; version 4.5				
	Sequencing vector: plasmid; L08752; 100% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Consensus quality: 184692 bases at least Q40				
	Consensus quality: 185645 bases at least Q30				
	Consensus quality: 186064 bases at least Q20				
	Insert size: 186479; sum-of-contigs				
	Insert size: 117845; 70.8% error; agarose-gel				
	Quality coverage: 5.87x in Q20 bases; sum-of-contigs				
	Quality coverage: 10.17x in Q20 bases; agarose-gel				
	NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence				

Suita, Osaka 565, Japan (Tel:81-6-6879-5111(ex.3211),
Fax:81-6-6879-3219)

FEATURES
source location/Qualifiers

1..1277
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XPACX1"
1..1277
/gene="XPAC"
1..67
/gene="XPAC"
68..871
/gene="XPAC"
/note="Xeroderma Pigmentosu Group A complementing protein"
/citation=[1]
/codon_start=1
/evidence=experimental
/product="xpac1 protein"
/protein_id="BAA06692.1"
/db_xref="GI:505148"
/translation="MEPEPEPEQENKEKEKLSAARAKTERNRQRLMRLQRLAC
RPPYPTGEGISTVAPKPPVDSGGFFIEEENAEQHVENVAPQPVLECDYLIEEC
GKDFMDSYLSNHFDLAVDCSDRAEKKHLITRTAEKQEYLKDCDIDKREPLKFIIL
KKNPHNTGMDKMYLKAQVIRKSLVWGSSEALBEAKVRKNDKMKQKKFDKKVK
ELRRVRSLSMKKEASGHQHEYGPEHVEDSYKKTCTTCGYENMYEKK"
872..1277
/gene="XPAC"
894..899
/gene="XPAC"
1257..1262
/gene="XPAC"
1277
/gene="XPAC"

3'UTR
polyA_signal
polyA_signal
polyA_site
ORIGIN
Query Match
Best Local Similarity 92.0%; Score 18.4; DB 5; Length 1277;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTATG 20
|||||
DB 792 GGTCCATCTCATGTGTATG 773

RESULT 14
AL732555 157653 bp DNA linear ROD 25-JUL-2002
LOCUS AL732555/c
DEFINITION Mouse DNA sequence from clone RP23-124119 on chromosome 4, complete
sequence.
ACCESSION AL732555
VERSION AL732555.7 GI:21998253
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 157653)
Lovel, J.
Direct Submission
Submitted (25-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
On Jul 26, 2002 this sequence version replaced gi:21955637.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk

COMMENT
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WormPep; information on the WormPep
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-124119 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Peter de Jong.
For further details see http://www.choiti.org/bacpac/home.htm
VECTOR: pBAC3.6.

FEATURES
source location/Qualifiers
1..157653
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-124119"
/clone_11b="RPCI-23"

ORIGIN
Query Match
Best Local Similarity 92.0%; Score 18.4; DB 10; Length 157653;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTATG 20
|||||
DB 143430 GGTCCATCTCATGTGTATG 143411

RESULT 15
AC109694 199263 bp DNA linear HTG 22-SEP-2002
LOCUS AC109694
DEFINITION Rattus norvegicus clone CH230-200G12. *** SEQUENCING IN PROGRESS
AC109694
VERSION AC109694.5 GI:2195474
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 199263)
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Angilano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Cavazos, J., Ceeasar, H., Center, A.,
Cardenas, V., Carter, K., Cavazos, J., Chen, Y., Chen, Z., Chu, J.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleaveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Dayla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Guerra, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunaratne, P., Haaland, N., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpethy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kows, C., Kratz, C.L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenschewski, L., Louised, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martineau, E., Mawani, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milsavayevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunolu, G., Olamunsgoon, A., Pal, S., Parks, K., Pasernack, S., Paul, H., Perez, A., Perez, L., Plankoch, C., Plopper, F., Polindexter, A., Popovic, D., Pritts, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, R., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, T., Thomas, N., Thomas, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wood, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 199263)
Worley, K.C.

Submitted (07-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 199263)
Rat Genome Sequencing Consortium.

Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 19, 2002 this sequence version replaced gi:22953896.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GGBN
Center clone name: CH230-200G12

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 139788 bases at least Q30
Consensus quality: 144250 bases at least Q30
Consensus quality: 147546 bases at least Q20
Estimated insert size: 170183; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 184185: contig of 184185 bp in length
* 184186 184285: gap of unknown length
* 184286 194997: contig of 10712 bp in length
* 194998 195097: gap of unknown length
* 195098 196262: contig of 1165 bp in length
* 196263 196362: gap of unknown length
* 196363 197533: contig of 1171 bp in length
* 197534 197633: gap of unknown length
* 197634 199263: contig of 1630 bp in length.

Location/Qualifiers
1. 199263
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-200G12"
1. 1119
/note="wgs end-extension
clone_end=T"
1129. 9971
/note="clone boundary
clone_end=T"
site:ECORI
end sequence: BH300659"
23827. 24368
/note="clone boundary
clone_end:Sp6
site:ECORI
end sequence: BH300660"
33304. 36846
/note="wgs end-extension
clone_end:Sp6"

misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Query Match
Best Local Similarity 95.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0;
Gaps 0;

1 168029 GGTCCATCTACTATGTTGATG 20
|||||
Db 168029 GGTCCATCTACTATGTTGATG 168048

RESULT 16
AC126180
LOCUS
DEFINITION
AC126180
AC126180.4 GI:30581548
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE
Rattus norvegicus
Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 227524)
Murny, D., Maric, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Bernabed, F., Biwaio, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, D., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guenara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harrey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howell,S., Huily,K.S., Hume,J., Idlebit,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpaty,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Louisedge,H., Lozano,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawany,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundaas,M., Murphy,M., Nair,L., Nankervils,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackeleme,O., Okunolu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Rutz,S.U., Sanders,M., Savary,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Soza,J., Steidle,M., Strong,R., Sutton,A., Svatek,A., Tabors,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K., Vales,R., Vera,V., Villaseca,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Woden,H., Woley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yan,D., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 227524)
Worley,K.C.
Direct Submission
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 227524)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:24942467.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKNS
Center clone name: CH230-208J22
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 194004 bases at least Q40

Consensus quality: 197089 bases at least Q30
Consensus quality: 199168 bases at least Q20
Estimated insert size: 205215; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 227524 contig of 227524 bp in length.
Location/Qualifiers
1..227524
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-208J22"
1..2135
/note="wgs_end_extension
clone_end:"77"
4585..5409
/note="clone_boundary
clone_end:"77"
site:ECORI
end_sequence:RMBKP59TJB"
226063..226880
/note="clone boundary
clone_end:Sp6
site:ECORI
end_sequence:RMBKP59TJB"

FEATURES

source

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 227524;
Best Local Similarity 95.0%; Pred. No. 1.5e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20
|||||
Db 58568 GGTCCTACTCATGTTGATG 58587

RESULT 17
AP003607/c 158826 bp DNA linear PLN 27-NOV-2003
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone: P0004A09.

ACCESSION
AP003607 BA000010
VERSION
AP003607.3 GI:15290074

KEYWORDS

SOURCE

ORGANISM

REFERENCE
AUTHORS

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
Katayose,Y., Wu,J., Nilmura,Y., Cheng,Z., Nagamura,Y.,
Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arikawa,K.,
Chien,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K.,
Hamada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y.,
Idonuma,A., Iijima,M., Ikeda,M., Ikeno,M., Itoh,S., Itoh,T.,
Itoh,Y., Iwabuuchi,A., Kamiya,K., Karsawa,W., Katagiri,S.,
Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara,T.,
Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
Ohta,I., Ono,N., Saij,S., Sakai,K., Shibata,M., Shimokawa,T.,
Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,

TITLE The genome sequence and structure of rice chromosome 1
JOURNAL Nature 420 (6913), 312-316 (2002)

CDS
39061..40536
/gene="P0004A09.9"
/note="contains ESTs C26603 (C12674), AU075575 (C12674)"
/codon_start=1
/product="putative salicylate-induced glucosyltransferase"
/protein_id="BAB63773.1"
/db_xref="GI:15290080"
/translation="NAIKDEOQPLHILFPPFLAPGLHPIADMALFARGVCTILT
TPNAAVIRSAVDRANDSPRRNGLAIETVPPDVLGPPGPGSGTALTODDRDK
PLGIRLHEPDRILSEHNDAAYVDSFERAAADAAHEGPGIGTCTSVFARACT
NSMLKNPLHETAPDDPAVPLPLGLPHCVELKRSQMDPKKPDHMEKROSIDADOK
SFEVFNSEFHELPDYVEHYRTTIGRRVWLVANLKDVAVRSSTLSLSPADGYLR
MIDAKPRGSVVVSVFGLTSSFSFPAEMRELARGLDLSGKFWVINGADADASBEMPEG
FAELIAPRGRELITIRGAPOMLILNHPAVGGTGVCGNNTLEAVTAGVPMVTPRY
ADQFNKELITVLELVGVGVSMDPASKLENRRVITGGVVGALGRVWGDEEBEAI
RKKATELGKAGALEKGGSSYDVGILMDLMARGSNNV"
complement(join(43936..43947,44152..44568))
/gene="P0004A09.10"
complement(join(43936..43947,44152..44568))
/gene="P0004A09.10"
complement(join(43936..43947,44152..44568))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB63774.1"
/db_xref="GI:15290081"
/translation="NMTDAARRARRPRTMWSGAPAVEAEAAAMIRSGGGGGGGRG
RLDDGMDLLEGVDNDDSTTAARQWDLLEGVDNDDSTTAARRRPRLSSPPMPSVVLA
SETATDIPPEPSHSSAHPLASPLPHERLDSVAPILA"
complement(join(45670..46125,46162..46395))
/gene="P0004A09.11"
complement(join(45670..46125,46162..46395))
/gene="P0004A09.11"
complement(join(45670..46125,46162..46395))

Query Match
Best Local Similarity 94.7%; Pred. No. 5e+02; Length 158826;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGAT 19
|||||
Db 25634 GGTCCATCTCATGTTGAT 25616
|||||

RESULT 18
AP004367/c 197674 bp DNA linear HTG 27-NOV-2003
LOCUS Oryza sativa (japonica cultivar-group) chromosome 1 clone P0696E01,
DEFINITION *** SEQUENCING IN PROGRESS ***.
ACCESSION AP004367 BM000010
VERSION AP004367.2 GI:21952912
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
1 Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
Katayose,Y., Wu,J., Nilmura,Y., Cheng,Z., Nagamura,Y.,
Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,S., Arikawa,K.,
Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arica,K.,
Hamada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y.,
Idonuma,A., Iijima,M., Ikeda,M., Ikono,M., Itoh,S., Itoh,T.,
Itoh,Y., Itoh,Y., Iwabuchi,A., Kamuya,K., Karasawa,M., Katagiri,S.,
Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Mebara,T.,
Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,
Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,
Yamagata,H., Yamane,H., Yoshiki,S., Yoshihara,R., Yukawa,K.,
Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,
Yano,M., Jiang,J. and Gojobori,T.
The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)
22337376
PUBMED 12447438

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
2 (bases 1 to 197674)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (19-NOV-2001) Takuji Sasaki, National Institute of
Agricultural Sciences, Rice Genome Research Program, Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tasasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jul 24, 2002 this sequence version replaced gi:17046147.
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 197674: contig of 197674 bp in length.
Location/Qualifiers
1..197674
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0696E01"

ORIGIN
Query Match
Best Local Similarity 94.7%; Pred. No. 5.1e+02; Length 197674;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGAT 19
|||||
Db 41502 GGTCCATCTCATGTTGAT 41484
|||||

RESULT 19
AC115118/c 198448 bp DNA linear ROD 05-NOV-2003
LOCUS Mus musculus BAC clone RP23-84N24 from 5, complete sequence.
DEFINITION AC115118
ACCESSION AC115118.3 GI:22138690
VERSION AC115118.3 GI:22138690
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198448)
The sequence of Mus musculus BAC clone RP23-84N24
Unpublished (2001)
2 (bases 1 to 198448)
Wilson,R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 198448)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (14-MAR-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 198448)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (28-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 198448)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 198448)
REFERENCE Wilson,R.
AUTHORS Direct Submission
TITLE Submitted (05-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL On Aug 8, 2002 this sequence version replaced gi:21218509.

Genome Center
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu

Summary Statistics
Center project name: M_BA0084N24

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-23 BAC library has been constructed by Kazutoyo Oseawa
and Minako Tateno in the laboratory of Pieter de Jong
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC122492:
Location/Qualifiers
1..198448
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-84N24"
/clone_1lb="RPCI-23"
repeat_region
673..857
/rpt_family="B2"
repeat_region
1952..2649
/rpt_family="L1"
repeat_region
4688..5150
/rpt_family="L1"
repeat_region
5186..5680
/rpt_family="L1"
repeat_region
6562..7050
/rpt_family="L1"
repeat_region
9091..9435
/rpt_family="MALR"
repeat_region
10463..10982
/rpt_family="ERVK"

repeat_region 11751..12096
/rpt_family="MALR"
repeat_region 1411..14209
/rpt_family="L1"
repeat_region 14210..14435
/rpt_family="L1"
repeat_region 14436..14620
/rpt_family="B2"
tRNA complement(14542..14615)
/product="tRNA-Ser"
/note="Likely pseudogene (MM Sc=42.36 / Sec struct
Sc=-22.25)"
repeat_region 14621..14951
/rpt_family="L1"
repeat_region 15355..16585
/rpt_family="L1"
repeat_region 16584..16622
/rpt_family="L1"
repeat_region 16623..17018
/rpt_family="MALR"
repeat_region 17019..18064
/rpt_family="L1"
repeat_region 18061..18488
/rpt_family="L1"
repeat_region 18477..18737
/rpt_family="L1"
repeat_region 18737..19336
/rpt_family="L1"
repeat_region 19721..20369
/rpt_family="L1"
repeat_region 21192..21376
/rpt_family="L1"
repeat_region 21377..21803
/rpt_family="ERVK"
repeat_region 21804..22035
/rpt_family="L1"
repeat_region 22058..22317
/rpt_family="L1"
repeat_region 22354..23361
/rpt_family="L1"
repeat_region 23798..24434
/rpt_family="L1"
repeat_region 26562..26741
/rpt_family="B2"
repeat_region 26886..27097
/rpt_family="B2"
repeat_region 27372..27501
/rpt_family="B2"
repeat_region 27681..27791
/rpt_family="A1u"
repeat_region 27806..27857
/rpt_family="MER1B"
repeat_region 29422..29579
/rpt_family="B4"
repeat_region 30476..30932
/rpt_family="L1"
repeat_region 30931..31525
/rpt_family="L1"
repeat_region 31526..31857
/rpt_family="L1"
repeat_region 31901..32233
/rpt_family="MALR"
repeat_region 35194..36055
/rpt_family="ERVK"
repeat_region 36058..36391
/rpt_family="ERVK"
repeat_region 36405..36491
/rpt_family="B4"
repeat_region 37726..38070
/rpt_family="MALR"
repeat_region 38166..38305
/rpt_family="MALR"
repeat_region 38306..38730

```
repeat_region      /rpt_family="ERVX"
38731. .38846
/rpt_family="MaLR"
repeat_region      39695. .39842
/rpt_family="Alu"
repeat_region      42175. .42378
/rpt_family="MER1_type"
repeat_region      42583. .42754
/rpt_family="B2"
repeat_region      42758. .42962
/rpt_family="B2"
repeat_region      43332. .43672
/rpt_family="L1"
repeat_region      44595. .44811
/rpt_family="B2"
repeat_region      45045. .45215
/rpt_family="B2"
repeat_region      45267. .45414
/rpt_family="Alu"
repeat_region      45570. .45777
/rpt_family="MaLR"
repeat_region      47156. .47217
/rpt_family="B2"
repeat_region      47218. .47556
/rpt_family="MaLR"
repeat_region      47557. .47670
/rpt_family="B2"
```

Query Match	87.0%;	Score 17.4;	DB 10;	Length 198448;
Best Local Similarity	94.7%;	Pred. No. 5.1e+02;		
Matches 18; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	2	GTCCATCTCATGTTGATG	20
Db	102771	GTCCATGCTCATGTTGATG	102753

RESULT 20	
LOCUS	HUAC004682
DEFINITION	Homo sapiens Chromosome 16 BAC clone CIT987SK-A-259H10, complete sequence.
ACCESSION	HUAC004682
VERSION	AC004682.1
KEYWORDS	GT:3337388
SOURCE	HTG.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 189134) Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R., Fuhrmann,J., Mason,T., Crosby,M.L., Barnstead,W., Cronin,L., Desai,R., May,A., Cao,Y., Xu,R.X., Kang,H.L., Mitchell,S., Eichler,E.E., Harris,P.C., Venter,J.C. and Adams,M.D. Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q Genomics 60 (3), 295-308 (1999)
JOURNAL	
MEDLINE	99425270
PUBMED	10493829
REFERENCE	2 (bases 1 to 189134) Adams,M.D., Loftus,B.J., Zhou,L., Crosby,M., Fuhrmann,J., Mason,T.M., Brandon,R., Kim,U.J., Kerlavage,A.R. and Venter,J.C. Homo sapiens Chromosome 16 BAC clone CIT987SK-A-259H10 Unpublished
AUTHORS	3 (bases 1 to 189134) Adams,M.D. and Loftus,B.J.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, Email: bjlloftus@tigr.org
REFERENCE	4 (bases 1 to 189134) Adams,M.D.
AUTHORS	Direct Submission
TITLE	

JOURNAL COMMENT
Submitted (24-JUL-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
On Jul 24, 1998 this sequence version replaced gi:3249114.

Submitted 44-JUL-1996) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
On Jul 24, 1998 this sequence version replaced gi:3249114.
Addresses all correspondence to: Mark Adams The Institute for Genomic Research 9712 Medical Center Dr. Rockville, MD 20850, USA e-mail address: humgen@tigr.org. The orientation of the sequence is from sp6 end to 3' end. Genes were identified by a combination of five methods including: XGRail (available by anonymous ftp from arthur.gem.ornl.gov/), Genefinder (Phil Green, University of Washington), GenScan (Chris Burge, <http://genome.stanford.edu/~chris/GENSCANv.html>) searches of the complete sequence against a peptide database, and the Human gene index database at TIGR (<http://www.tigr.org/cdb/hgi.html>). Genes without peptide homology having spliced EST hits are termed 'Unknown gene product'. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).

FEATURES

gene
mRNA

Gene	/clone="A-259H10.1"	/complement{133825..>150585}
mrna	/gene="A-259H10.1"	
	complement{10318,133825..14330,135666..135787,137233..137328,137818..137943,138136,138270,138361..138516,138944..139098,139195..139403,140626..140738,141119..141226,141395..141511,141606..141716,142117..142252,142612..142779,142955..143141,143532..143669,145131..145244,145550..145657,146146..146307,146856,147011,147463..147533,147698..147805,147931,148496,149066,149793..149921,150263..150585}	
	/gene="A-259H10.1"	
SFS	133855..133981	

STS

CDS

/note="J39391, EST155365, Chr. 16, Homo sapiens"
 /db xref="dbSTS:G25538"
 13855..13381
 /note="A03992, SHGC-32168, Chr. 16, Homo sapiens"
 /db xref="dbSTS:G27137"
 Complement(join(1338246..134330,135666..135767,
 137233..137328,137818..137944,138136..138270,
 138361..138516,138944..139098,139195..139403,
 140626..140739,141119..141226,141365..141571,
 141606..141716,142117..142252,142612..142787,
 142955..143141,143532..143669,145311..145243,
 143550..144857,146146..146307,146856..147011,
 147463..147539,147698..147816,147951..148098,
 148962..149066,149734..149921,150263).150585)
 (gene="PA-359H0.1")

```

/codon_start=1
/product="pre-mRNA splicing factor (PR16) (K1AA0224)"
/protein_id="PAC27431.1"
/db_xref="01:3373189"
/translation="MCGTSSDAAIHRLGGTDLCQVGLICKSKAASEQHFKAAPAF
RPSLLGGLSLAVLPRERKEDEGDEGDEGCKKYSVYKDMWESDQADCKDEEGSGQAG
NIRKORHYSALVSTPPIPGVGEEMERSQERPMDEKSYAASEKKEKQKQKQKQKQKQ
DRDYKSRKQWESRPSPTPRSDSGESSRNSRPSRRPDAQTPRSTWEESESESESESE
DGSYSKSRQWESRPSPTPRSDSGESSRNSRPSRRPDAQTPRSTWEESESESESESE
MADRRHRLGSTRPSRSGRGESESESESPHPESESESESESESESESESESESESESE
FEHNLALASDYLVRKREQLHKQKQKRIISAQRQINENEMENRNLTSGVHRLLE
VDEDEENAKNAVLVHNLVPPFLDGRVLTQKPEPVLVDAKASDIALIARKSQD
VAKHNEQSEKRAQKHMKELAGKGLDINGVKKEEEDPAVDEKRVYRDEQKADQ
MKKSEASEEKAKKSLTBORQVLPFAVQKELLIRNNSVIVGGTGGKQTQL
IKYMDGDTLDEGKQHPYRPAVAMSAKRVSEMGNGVGEVAFIBEDTSNTIT
TSAMDAKPFAFGVNPPIHIGRFPVPLISCTPQEDYVAVAKVOSLQVLSGAL
GDLILFMWGQDLEVSQDVLPIHLEELANALPLVPIYSLPSLQKIRPKADG
RKCIVAINIATSLTVDGIVNFVYDSCTCKATFNPRIGMDLQIYPIQSANANSG
AATRGPGCCFLYQVSAKBLTGLTVSTPELQRLNLAVALKRSAGVQDLQPFEMD
PEDMNNMFLQMLIALDNLNTLSTGLVLEFPLDALSKMLVSDMGCSILL

```


LIIVSLVPAIFRPRKREESDOIKEKFAVESDHLITVINYLOKNNSTIWCND
HFIHAKAMKREVEVAQOLKDIWOOOBSIASGOTPMIDIVKICCAVPHQAKLGIG
EYVNIIRITGMPCHLIPSSIFGMYTPDYIVYHEIWTTEWOCYTAANDFELIATIG
MEYSYKQAKSKQENRRRAKERAASAMEEMALAEOLRARROEQRKSPVLSVSTKI
YTPGKEQGEPTPRTPARFGL"
complement(<169670..>172858)
/gene="A-259H10.2"
complement(join(169670..170448,171532..171606,
172366..172467,172774..>172858))
/gene="A-259H10.2"
complement(join(169670..170448,171532..171606,
172366..172467,172774..>172858))
/gene="A-259H10.2"
/codon_start=1
/product="haploglobin-related protein precursor"
/protein_id="AAC27433.1"
/db_xref="GI:3337391"
/translation="DLGAVYSLILNGRQLFALYSGNDVTDISDRFPKPEIANGYVE
HLFRYCKNYRLRTGEGVYTLNDKQWINKAVGDKLPECAVCGKPKNPANPYQRI
LGGHLDKAGSPFMOAKMVSHNLTTGATLINQOMLITTAKNLFLNHSNATAKDIAPT
LTLVYKQVLVEIEKYVLAHPNHOVDIGIKLKOKYLVNERMPICLBKRYAARV
GVSGMGSDNFKLTDHLKYVLPVADQYDCTTHTEGSTCPKPKAPKSPVQPIINE
HFCVGMSTYQEDTCYGDAGSAFAVHDLDEDTWYAAILSPDKSCAVAEYGVYKVT
IOHWYQXTIAEN"
175654..175784
/note="35342..WT-17126..Chr. 16, Homo sapiens"
/db_xref="dbSTS:G21504"
complement(<185861..188871)
/gene="A-259H10.3"
complement(join(<185861..186639,187563..187637,
188396..188497,188784..188871))
/gene="A-259H10.3"
complement(join(185861..186639,187563..187637,
188396..188497,188784..>188865))
/gene="A-259H10.3"
/codon_start=1
/product="haploglobin"
/protein_id="AAC27432.1"
/db_xref="GI:3337390"
/translation="ALGAVIALLLNGQLFVAVDSNDVTDIADGCPKPEIANGYVEH
SVRYCKNYKRLRTGEGVYTLNNEKQWINKAVGDKLPECAVCGKPKNPANPYQRI
LGGHLDKAGSPFMOAKMVSHNLTTGATLINQOMLITTAKNLFLNHSNATAKDIAPT
LTLVYKQVLVEIEKYVLAHPNHOVDIGIKLKOKYLVNERMPICLBKRYAARV
GVSGMGSDNFKLTDHLKYVLPVADQYDCTTHTEGSTCPKPKAPKSPVQPIINE
HFCVGMSTYQEDTCYGDAGSAFAVHDLDEDTWYAAILSPDKSCAVAEYGVYKVT
QHWYQXTIAEN"

ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 189134;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CCATACCTCATGTTGATG 20
|||
Db 104856 CCATACCTCATGTTGATG 104872

RESULT 21
AC009160/c
LOCUS AC009160.5 GI:8575975
DEFINITION 27 unordered pieces.
AC009160 190856 bp DNA linear HTG 21-JUN-2000
Homosapiens chromosome 16 clone RP11-70E3, WORKING DRAFT SEQUENCE,
AC009160
VERSION AC009160.5 GI:8575975
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 190856)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

2 (bases 1 to 190856)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jun 21, 2000 this sequence version replaced gi:7689973.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 433344
Center clone name: RPCI-11_70E3

Summary Statistics
Consensus quality: 162008 bases at least Q40
Consensus quality: 174412 bases at least Q30
Consensus quality: 177602 bases at least Q20
Estimated insert size: 162600; agarose-ep estimation
Estimated insert size: 188256; sum-of-contigs estimation
Quality coverage: 5.83 in Q20 bases; agarose-ep estimation
Quality coverage: 5.04 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1014: contig of 1014 bp in length
* 1015 1114: gap of unknown length
* 1115 2667: contig of 1553 bp in length
* 2668 4750: contig of 1983 bp in length
* 4751 4851: gap of unknown length
* 4851 6226: contig of 1376 bp in length
* 6227 6326: gap of unknown length
* 6327 7638: contig of 1312 bp in length
* 7639 7739: gap of unknown length
* 7739 8805: contig of 1067 bp in length
* 8806 8905: gap of unknown length
* 8906 10073: contig of 1168 bp in length
* 10074 10173: gap of unknown length
* 10174 11248: contig of 1075 bp in length
* 11249 11349: gap of unknown length
* 11349 12369: contig of 1020 bp in length
* 12369 12468: gap of unknown length
* 12469 13733: contig of 1265 bp in length
* 13734 13833: gap of unknown length
* 13834 15022: contig of 1189 bp in length
* 15023 15122: gap of unknown length
* 15123 16567: contig of 1445 bp in length
* 16568 16667: gap of unknown length
* 16668 18356: contig of 1689 bp in length
* 18357 18456: gap of unknown length
* 18457 22179: contig of 3723 bp in length
* 22180 22279: gap of unknown length
* 22280 25036: contig of 2757 bp in length
* 25037 25137: gap of unknown length
* 25137 31395: contig of 6259 bp in length
* 31396 31495: gap of unknown length
* 31496 39492: contig of 7997 bp in length
* 39493 39593: gap of unknown length
* 39593 48267: contig of 8674 bp in length
* 48267 48367: gap of unknown length
* 48367 60754: contig of 12388 bp in length
* 60755 60854: gap of unknown length
* 60855 71400: contig of 10546 bp in length
* 71401 71500: gap of unknown length
* 71501 81265: contig of 9765 bp in length
* 81266 81365: gap of unknown length

```

*      81366      95666: contig of 14301 bp in length
*      95767      108787: gap of unknown length
*      95767      108787: contig of 13021 bp in length
*      108788      127457: gap of unknown length
*      108888      127457: contig of 18569 bp in length
*      127457      127556: gap of unknown length
*      127556      147174: contig of 19618 bp in length
*      147175      147274: gap of unknown length
*      147275      168580: contig of 21306 bp in length
*      168581      168680: gap of unknown length
*      168681      190856: contig of 22176 bp in length.
FEATURES
  source
    1..190856
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="16"
    /clone="RP11-70E3"
    /clone_1ib="RPci human BAC library 11"
ORIGIN
Query Match      85.0%; Score 17; DB 2; Length 190856;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      4 CCATCTCATGTTGATG 20
Db      157563 CCATCTCATGTTGATG 157547
RESULT 22
AC134076      253297 bp      DNA      linear      HTG 13-NOV-2002
LOCUS      Rattus norvegicus clone CH230-2F4, WORKING DRAFT SEQUENCE.
ACCESSION      AC134076.2 GI:24941326
VERSION      HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS      Rattus norvegicus
SOURCE      Rattus norvegicus
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
  1 (bases 1 to 253297)
  Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
  Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
  Barbarta,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
  Bouck,U., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
  Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
  Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
  Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
  Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
  Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
  Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
  Douthwaite,K.U., Draper,H., Dugan-Rocha,S., Durbin,K.U.,
  Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
  Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
  Gabrieli,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
  Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
  Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
  Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
  Homel,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.B.,
  Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
  Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
  Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
  Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
  Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
  Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
  Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
  Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
  Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
  Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguni,M., Okunolu,G.,
  Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
  Peters,L., Pickens,R., Plimms,E., Pu,L.L., Quiles,M., Ren,Y.,

```

```

Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoochattai,N., Sisson,I.,
Sodergren,E., Sotniko,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swalek,A., Tabor,P., Tamerias,A., Tamerias,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,J., Vera,V., Villalton,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 253297)
Worley,K.C.
Direct Submission
Submitted (22-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253297)
Worley,K.C.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23267443.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TURX
Center clone name: CH230-2F4
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 234461 bases at least Q40
Consensus quality: 236497 bases at least Q30
Consensus quality: 237862 bases at least Q20
Estimated insert size: 242145; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 253297: contig of 253297 bp in length.
FEATURES
  source
    1..253297
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-2F4"
    /clone="CH230-2F4"
    /note="wgs_contig"
misc_feature

```



```

site:EcORI
end_sequence:BH309953"
31376. .33614
/misc_feature
/nc="wgs_contig"
40354. .44099
/misc_feature
/nc="wgs_contig"
123395. .124915
/misc_feature
/nc="wgs_contig"
179305. .180899
/misc_feature
/nc="wgs_contig"
210644. .211731
/misc_feature
/nc="wgs_contig"
211782. .213121
/misc_feature
/nc="wgs_contig"
246452. .247376
/misc_feature
/nc="clone_boundary
clone_end:T7
site:EcORI
end_sequence:BH309950"
249386. .250543
/misc_feature
/nc="wgs_end_extension
clone_end:T7"
260312. .262005
/misc_feature
/nc="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match 85.0%; Score 17; DB 2; Length 263730;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCCATACATCATGTTG 17
|||||
Db 167351 GGTCCATACATCATGTTG 167335

RESULT 24
MXXPAC6 294 bp DNA linear ROD 06-FEB-1997
LOCUS Xeroderma pigmentosum group A Correcting gene, exon
DEFINITION
6.
ACCESSION X74350 GI:440563
VERSION X74350.1
KEYWORDS Pigmentosum group A Correcting protein (XPAC).
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 294)
van Oostrom,C.T., de Vries,A., Verbeek,S.J., van Kreijl,C.F. and
van Steeg,H.
Cloning and characterization of the mouse XPAC gene
Nucleic Acids Res. 22 (1), 11-14 (1994)
94173654
8127648
2 (bases 1 to 294)
van Steeg,H.
Direct Submision
Submitted (27-JUL-1993) H. Van Steeg, National Inst of Public
Health and Environmental Protection, P O B 1, 3720 BA Bilthoven,
NETHERLANDS
FEATURES
Location/Qualifiers
1..294
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="Ola129"
/db_xref="taxon:10090"
/chromosome="4C2"
39. .>215
/usedin=X74345:XPAC_CDS
/label=ex6
210. .215
polyA_signal

```

[illegible]

```

AUTHORS
van Oostrom,C.T., de Vries,A., Verbeek,S.J., van Kreijl,C.F. and
van Steeg,H.
Cloning and characterization of the mouse XPAC gene
JOURNAL
Nucleic Acids Res. 22 (1), 11-14 (1994)
MEDLINE
94173654
REFERENCE
PUBMED
2 (bases 1 to 938)
AUTHORS
van Steeg,H.
TITLE
Direct Submission
JOURNAL
Submitted (27-JUL-1993) H. Van Steeg, National Inst of Public
Health and, Environmental Protection, P O B 1, 3720 BA Bilthoven,
NETHERLANDS

FEATURES
source
location/Qualifiers
1..938
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Ola 129"
/db_xref="taxon:10090"
/cell_line="embryonic stem cell line E14"
1..938
/gene="XPAC Xeroderma Pigmentosum group A Correcting gene"
92..910
/gene="XPAC Xeroderma Pigmentosum group A Correcting gene"
/codon_start=1
/product="XPAC protein"
/protein_id="CAA52393.1"
/db_xref="GI:440565"
/db_xref="GOA:O64267"
/db_xref="SWISS-PROT:O64267"
/translation="MATHAEKQTSPEEPALADPEAPLPAAYASVERKORALMROAR
LAAPRPAAATGVAASVKAPKMDYGGFILLEEKKEHIGNIVHGPVMEVDYTA
ICEBGKEFMDISLMNHFDLPDCSCRPADFKHLITTEAKOEYLKDCLEKPEAPK
LRFLVKNPRIHRSOWGDMKLYLKLGVKKALFVMSQEALIDAKERQENREKMKOKK
DKVKELEBARLRSSVWKRETTTHQHEVGPENLEDMDVRYKTCILGHETVKKM"
933..938
/gene="XPAC Xeroderma Pigmentosum group A Correcting gene"

ORIGIN
Query Match 84.0%; Score 16.8; DB 10; Length 938;
Best Local Similarity 90.0%; Pred. No. 7.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
Db 831 GGTCCATCTCATGTTGATG 812

RESULT 27
AX566974 1029 bp DNA linear PAT 29-NOV-2002
LOCUS AX566974
DEFINITION Sequence 173 from Patent WO02077021.
ACCESSION AX566974
VERSION AX566974.1 GI:26001549
KEYWORDS
SOURCE
ORGANISM
Streptococcus pneumoniae
Streptococcus pneumoniae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE
1
Waeigman,V., Tettelin,H. and Fraser,C.
Streptococcus pneumoniae proteins and nucleic acids
Patent: WO 02077021-A 173 03-OCT-2002;
Chiron Spa (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES
source
location/Qualifiers
1..1029
/organism="Streptococcus pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:1313"

ORIGIN
Query Match 84.0%; Score 16.8; DB 6; Length 1029;
Best Local Similarity 90.0%; Pred. No. 7.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

RESULT 28	AKI15474/c	2138 bp	mRNA	linear	INV 30-NOV-2002
LOCUS	AKI15474				
DEFINITION	Ciona intestinalis cDNA, clone:cig055b24, full insert sequence.				
ACCESSION	AKI15474				
VERSION	AKI15474.1				
KEYWORDS	FLI cDNA.				
SOURCE	Ciona intestinalis				
ORGANISM	Ciona intestinalis				
REFERENCE	Bukacinski, Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.				
AUTHORS	1 Satou, Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T., Sasaki, A., Hamaguchi, M., Awazu, S., Yagi, K., Sasakura, Y., Nakayama, A., Ishikawa, H., Inaba, K. and Satoh, N.				
TITLE	A cDNA resource from the basal chordate Ciona intestinalis				
JOURNAL	Genesis 33 (4), 153-154 (2002)				
MEDLINE	22191024				
PUBMED	12203911				
REFERENCE	2 (bases 1 to 2138)				
AUTHORS	Satou, Y. and Satoh, N.				
TITLE	Direct Submission				
JOURNAL	Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of Zoology; Sakyo-ku, Kyoto, Kyoto 606-8502, Japan (E-mail:satoh@ascidian.zool.kyoto-u.ac.jp, Tel:81-75-753-4095, Fax:81-75-705-1113)				
COMMENT	Ciona intestinalis cDNA Project (URL: http://ghost.zool.kyoto-u.ac.jp/index1.html).				
FEATURES	Location/Qualifiers				
SOURCE	1..2138				
	/organism="Ciona intestinalis"				
	/mol_type="mRNA"				
	/db_xref="taxon:7719"				
	/clone="cig055b24"				
ORIGIN					
Query Match	84.0%; Score 16.8; DB 3; Length 2138;				
Best Local Similarity	90.0%; Pred. No. 8.2e+02;				
Matches	18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
OR	1 GGTCATCTCATGTTGATG 20				
DB	1185 GGTCATCTCATGTTGATG 1166				
RESULT 29	AE007333/c	11344 bp	DNA	linear	BCT 31-AUG-2001
LOCUS	AE007333/c				
DEFINITION	Streptococcus pneumoniae TIGR4 section 16 of 194 of the complete genome.				
ACCESSION	AE007333				
VERSION	AE007333.1				
KEYWORDS	GI:14971638				
SOURCE	Streptococcus pneumoniae TIGR4				
ORGANISM	Streptococcus pneumoniae TIGR4				
REFERENCE	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.				
AUTHORS	1 (bases 1 to 11344)				
	Tettelein, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Salzberg, S.L., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Ueberback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angilim, S., Dickinson, T., Hickey, E.K., Holt, I.E., Loftus, B.J., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.				

TITLE Complete genome sequence of a virulent isolate of Streptococcus pneumoniae
JOURNAL Science 293 (5529), 498-506 (2001)
MEDLINE 21357209
PUBMED 11463916
REFERENCE 2 (bases 1 to 13344)
AUTHORS Tetteelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., Deboy,R.T., Haft,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Lewis,M.R., Radune,D., Holtzapple,E., Knorr,H., Wolf,A.M., Uetebach,T.R., Hansen,C.J., McDonald,J.A., Felblyum,T.V., Angiuoli,S., Gesuwan,P., Hickey,E.K., Holt,I.E., Loftus,B.J., Uwal,M.L., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1. .11344
/organism="Streptococcus pneumoniae TIGR4"
/mol_type="genomic DNA"
/strain="TIGR4"
/db_xref="taxon:170187"
90. .194
/gene="SP0172"
90. .194
/gene="SP0172"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK74353.1"
/db_xref="GI:14971639"
/translation="MIFRYTILLPLADFYVEGLSGYVELFCSP"
255. .2204
/gene="SP0173"
255. .2204
/gene="SP0173"
/codon_start=1
/transl_table=11
/product="DNA mismatch repair protein HexB"
/protein_id="AAK74354.1"
/db_xref="GI:14971640"
/translation="MSHITLPEMLANOJAGVIERPAAVVKELVENADAGSQIITIEERGLKKVOITONGHIADEVLAARRHATSKIQALFRITRLTGFPGALPSIASVSVTLTLTAVDASHGTXYARAGEVEEYIPATSPVGTQCYEDLFFNTPALAKYMSQQAELSHIIDIVNRILGAPHEISFSLISQKENTRTAGTQLRQAIAGIYGLVSAKMWIEENSDLPFEISGFVSLPELTANNNYISLFINGRYIVNPLNRAIADGFSKLWGRFPLAVIHIHIDPYLADVNHPTKQVRISKEKELMTLVSEAIANSLKQOTLIPD ALENLAKSTVNRREKVEOTILPLKENTLYYEKTEPERSPOTEAADVQLTEGGDLT LAKETLDLTTRKAPKAPBRKPRANYDQLDHPILDASIDKAYDKLEREASPEPLE FFGQNHGTYLFRQGRDGLYITIDHAAQEKYKEEYRESIGNDOSQQLIVYIEFPAADALRLKERMPLLEEVGVFLAEGVNOFLIEHPIWMAEEIESGTYEMCDMLLTKEVSIRKTYRAELIIMMSCRSIRANNRIDHNSARQLLYQLSCDNPYNCPHGRPYLVHF TKSDEMKERRIOENHTSLRELGY"
complement (2340. .2432)
/gene="SP0174"
complement (2340. .2432)
/gene="SP0174"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK74355.1"
/db_xref="GI:14971641"
/translation="WVKHNFVDVDTGKISKHKCFEITDKTDV"
complement (2529. .2996)
/gene="SP0175"
complement (2529. .2996)
/gene="SP0175"
/note="identified by match to PFM protein family HM PF00885"
/codon_start=1

gene
CDS
/transl_table=11
/product="riboflavin synthase, beta subunit"
/protein_id="AAK74356.1"
/db_xref="GI:14971642"
/translation="MNTYEGNIVANNIKIGIVARFNEFITSKLSGALDNLKRENVN EKDIEVAVWGAFFIPLASAKMAKSKYDAILICGAVIRNTHSYDVCEVSGKIQ ISLNSLPVMEGVLTTDTIQLAERAGTKAGNKSBCQAQAIEMVNLIRTLDA"
complement (2997. .4202)
/gene="SP0176"
complement (2997. .4202)
/gene="SP0176"
/note="identified by match to PFM protein family HM PF00926"
/codon_start=1
/transl_table=11
/product="3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II"
/protein_id="AAK74357.1"
/db_xref="GI:14971643"
/translation="MEYRKIOEALQKRLVVIDDKRENEGDLICGAQAATTEN VNFMATYAKGLICMPSESLANQMLSPVVENNTDNHKTAFVSIIDYKETTGISAE RGLTARMGVADIEITPSDFRRGHMFPLIAKKGVLERNGTEATVDLLKLAGLKECGL CCEIIMNDGKMRPTDILIOSKKNITPLITIKLOERYKVDOLVERVTVNMPTRYG NFKAISYIDKUNGEHHLALINGNIEDBANVLCRHSECTIGDVLGSRCCGCGPDKA MKMIVENGSGVLLYLRQEGRGILNKLKAHYLDQGDQDULDANLGLFBODLETHI GAKMKIDGLQSLHLTNNPDKXQLERKGYTSSRISIEIANPYDSFYLETKNRM GHILNMEK"
complement (4222. .4857)
/gene="SP0177"
complement (4222. .4857)
/gene="SP0177"
/note="identified by match to TIGR protein family HM TIGR00187"
/codon_start=1
/transl_table=11
/product="riboflavin synthase, alpha subunit"
/protein_id="AAK74358.1"
/db_xref="GI:14971644"
/translation="MFTGIIIEIGKVERIQDSHNCKLSIKASKILTDHIGDSIAVN GILCTVHFHQSFVVDVNETWERTALTLTKHSEVNLSEALSVNRLGHHVVTGHI DGTGKISIKKDDNNAVYQINTOKEIIDLIVERGSTIIDIGSLTAVAKSVNFSVSI PHTLEQTILSKQVGYSTVNLENDILGKTVQKLMNSKSELSKELLQNGF"
complement (4842. .5942)
/gene="SP0178"
complement (4842. .5942)
/gene="SP0178"
/note="identified by match to TIGR protein family HM TIGR00326"
/codon_start=1
/transl_table=11
/product="riboflavin biosynthesis protein RibD"
/protein_id="AAK74359.1"
/db_xref="GI:14971645"
/translation="WSDSKIKWIKAIKLAQKAGVYVNPVGVAYIVKDNHIIIGCYHE FFGGHAERNALKNCRESPVATLYVTLPECHGKTPCIDAIIDSGITVIVGSLD CNPIVSGKGVKILENNLQVTVGILENCLMISFRKYITQHPVPMKXAMSDMD IATKNOGKWTIEBERAGVHQLRHVSAIIVGVNTYODDPLTRCLEBGNKPIRII CDTHLRPTLSKIVKTNIDIKTYIATSSEDGNKKMLVONHGCETLSIKKKNNHIDLS LMQHGNQIISLVLEGSIMMWGALBQIYDELKITYAPRVPGSAAKFPVGGGISTL PNDATRLRPVAFSQIGNDYLLIESEYIIPCSOE"
6347. .6940
/gene="SP0179"
6347. .6940
/gene="SP0179"
/note="identified by match to PFM protein family HM PF01330"
/codon_start=1
/transl_table=11
/product="Holliday junction DNA helicase RuvA"
/protein_id="AAK74360.1"
/db_xref="GI:14971646"
/translation="WYAYLKGITIKITAKYIVLETTNGIGYILVHVPYVSGVQNGEA OIVYHQVVRDEAHLLYGFRSDEKCLFSLISVSGIIPVSAIAIIPAADNGLVQAI

```

gene
  LKNITLYLTKPKIKKTAQOMVLDLGGKVVVAGDDLPAKVAVQASAEQNEJSEAMEAN
  LALGYKATELKKIKKFFEGTDTTANENYIKSAIKMLVK"
CDS
  /gene="SP0180"
  /locus="7513
  6977..7513
  /gene="SP0180"
  /codon_start=1
  /transl_table=11
  /product="DNA-3-methyladenine glycosylase I"
  /protein_id="AAK74361.1"
  /db_xref="GI:14971647"
  /translation="MTNPLXYAHDEWGGQPLHDDQVLFELLCMETYOAGLSWETVLN
  KROAREVFNHSGYDIAVEMDTLEAMLENPAIRNPAKLFATPANKQAFLQCAEY
  GSPFAYLWSFEGKTVNDVDPYRQAPAKTPLSEKLANDLKRKGFPGVAVLSFLQ
  AAGLVDDHENDCEMGLK"
  7510..8187
  /gene="SP0181"
  7510..8187
  /gene="SP0181"
  /note="identified by Glimmer2; putative"
  /codon_start=1
  /transl_table=11
  /product="conserved hypothetical protein"
  /protein_id="AAK74362.1"
  /db_xref="GI:14971648"
  /translation="MMSNKKELIFATILYTVLFMPDGVKLLASIMPSAIAINLYVVV
  LALYSGPLFKRRLIQMKKEIRKTRKFFRGVUTGMLPALMTVVEPSEMLKQVGL
  DQGLNQNSTQSTFOEOPULIAVFCVIGPLVEELPFROVLNHYOERLSGLSTILIV
  GLVFLVTHMHSIALSEWIGAVGVLGGGLAFSLIYVEKENIYPLVHMLSNSLSLI
  LAISIVK"
  complement(8324..9355)
  /gene="SP0182"
  complement(8324..9355)
  /gene="SP0182"

gene
  CDS

Query Match
  Best Local Similarity 84.0%; Score 16.8; DB 1; Length 11344;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy
  1 GGTCATACATCATGTGATG 20
  ||||| ||||| ||||| |||||
  Db 8381 GGTCATGCTCATGTAGATG 8362

RESULT 30
  AR218808 14273 bp DNA linear PAT 25-SEP-2002
  LOCUS AR218808
  DEFINITION Sequence 40 from patent US 6420135.
  ACCESSION AR218808
  VERSION AR218808.1 GI:23319742
  KEYWORDS
  SOURCE
  ORGANISM
  Unclassified.
  1 (bases 1 to 14273)
  Kunsch,C.A., Choi,G.H., Dillon,P.S., Rosen,C.A., Barash,S.C.,
  Fannon,M.R. and Dougherty,B.A.
  Streptococcus pneumoniae polynucleotides and sequences
  Patent: US 6420135-A 40 16-JUL-2002;
  TITLE
  JOURNAL
  FEATURES
  source
  1..14273
  /organism="unknown"
  /mol_type="genomic DNA"

ORIGIN
  Query Match
  Best Local Similarity 84.0%; Score 16.8; DB 6; Length 14273;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy
  1 GGTCATACATCATGTGATG 20
  ||||| ||||| ||||| |||||
  Db 13385 GGTCATGCTCATGTAGATG 13366
```

```

RESULT 31
  BD003720/c 14273 bp DNA linear PAT 31-JAN-2002
  LOCUS BD003720
  DEFINITION Polynucleotide of Streptococcus pneumoniae and sequence.
  ACCESSION BD003720
  VERSION BD003720.1 GI:18631681
  KEYWORDS JP 2001501833-A/40.
  SOURCE
  ORGANISM
  Streptococcus pneumoniae
  unclassified
  unclassified
  unclassified
  1 (bases 1 to 14273)
  Kunsch,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Barash,S.C.,
  Fannon,M. and Dougherty,B.A.
  Polynucleotide of Streptococcus pneumoniae and sequence
  Patent: JP 2001501833-A 40 13-FEB-2001;
  TITLE
  JOURNAL
  COMMENT
  HUMAN GENOME SCIENCES INC
  OS
  UNIDENTIFIED
  PN JP 2001501833-A/40
  PD 13-FEB-2001
  PF 30-OCT-1997 JP 1998520718
  PR 31-OCT-1996 US 60/029960
  PI CHARLES A KUNSCH,GIL H CHOI,PATRICK J DILLON,CRAIG A ROSEN, PI
  STEVEN C BARASH,
  PI MICHAEL FANNON,BRIAN A DOUGHERTY
  PC C12N15/09,A01K67/027,C07K14/315,C07K16/12,C12N1/15,C12N1/19,
  PC C12N1/21,
  PC C12N5/10,C12P21/02,C12Q1/68,G06F17/30,C12N15/00,C12N5/00, PC
  G06F15/40
  CC Strandedness: Double;
  CC Topology: Linear;
  FH key
  FT source
  1..14273
  /organism="unidentified".
  location/Qualifiers
  1..14273
  /organism="unidentified"
  /mol_type="genomic DNA"
  /db_xref="taxon:32644"

ORIGIN
  Query Match
  Best Local Similarity 84.0%; Score 16.8; DB 6; Length 14273;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy
  1 GGTCATACATCATGTGATG 20
  ||||| ||||| ||||| |||||
  Db 13385 GGTCATGCTCATGTAGATG 13366

RESULT 32
  HS398A12 90169 bp DNA linear PRI 04-MAR-2003
  LOCUS HS398A12
  DEFINITION Human DNA sequence from clone Rp3-398A12 on chromosome 6p23-25.1,
  complete sequence.
  ACCESSION AL021332
  VERSION AL021332.1 GI:2804159
  KEYWORDS
  SOURCE
  ORGANISM
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
  1 (bases 1 to 90169)
  Williams,S.
  Direct Submission
  Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  ----- Genome Center
  Center: Wellcome Trust Sanger Institute
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: humquerry@sanger.ac.uk
```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone configs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>

RP3-396A12 is from the library RPCT-3 constructed by the group of Plietier de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

FEATURES

source

1..90169
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCTP704A12398"
/db_xref="taxon:9606"
/chromosome="6"
/map="p23-25.1"
/clone="RP3-396A12"
/clone_1ib="RPCT-3"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 90169;

Best Local Similarity 90.0%; Pred. No. 1e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20

DB 60420 GCTCCATCTCATGTTGATG 60439

RESULT 33

AC095009_1/c

WPCOMMENT

Sequence split into 4 fragments LOCUS AC095009 Accession AC095009

Fragment Name Begin End

AC095009_0 1 110000

AC095009_1 100001 210000

AC095009_2 200001 310000

AC095009_3 300001 395260

Continuation (2 of 4) of AC095009 from base 100001 (AC095009 Rattus norvegicus clone CH2

Query Match 84.0%; Score 16.8; DB 2; Length 110000;

Best Local Similarity 90.0%; Pred. No. 1e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20

DB 94388 GGCCTATCTCATGTTGATG 94369

RESULT 34

BX537312/c

LOCUS BX537312 121113 bp DNA linear HTG 24-SEP-2003

DEFINITION Danio rerio clone DKEX-6P9, WORKING DRAFT SEQUENCE, 4 unordered

ACCESSION BX537312
VERSION GI:35209488
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.

REFERENCE

McLaren, S.

Direct Submission

Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Sep 24, 2003 this sequence version replaced gi:33414383.

COMMENT

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zf1sh-help@sanger.ac.uk

Project Information

Center project name: zK6P9

Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 120006 bases at least Q40

Consensus quality: 120458 bases at least Q30

Insert size: 120813; sum-of-contigs

Insert size: 126278; 4.2% error; agarose-fp

Quality coverage: 11.33x in Q20 bases; sum-of-contigs Quality

coverage: 10.84x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 63519: contig of 63519 bp in length
* 63520 63619: gap of 100 bp
* 63620 88215: contig of 24596 bp in length
* 88216 88315: gap of 100 bp
* 88316 110692: contig of 22377 bp in length
* 110693 110792: gap of 100 bp
* 110793 121113: contig of 10321 bp in length.

FEATURES

source

1..121113
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_1ib="DKEX-6P9"
/clone_1ib="DanioKey"

1..63519
/note="assembly_fragment:00731
fragment_chain:1"

63620..88215
/note="assembly_fragment:01659
fragment_chain:1"

88316..110692
/note="assembly_fragment:00742.0"

110793..121113
/note="assembly_fragment:00871"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 121113;

Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2

QY 1 GGTCCATCTCATGTTGATG 20

Db 49662 GGTCATCTCATGTGATG 49643

RESULT 35
LOCUS AP003965/c 121852 bp DNA linear PRI 15-MAR-2003
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:CM9-3109, complete sequence.
ACCESSION AP003965
VERSION AP003965.2 GI:17426130
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2001)
AUTHORS 2 (bases 1 to 121852)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22, Suenho-chou, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@psc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On Dec 7, 2001 this sequence version replaced gi:15027546.

FEATURES
source location/Qualifiers
1..121852
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="CM9-3109"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 121852;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGATG 20
Db 64058 GGACCATCTCATGTGATG 64039

RESULT 36
LOCUS HS326L13/c 127247 bp DNA linear PRI 07-AUG-2003
DEFINITION Human DNA sequence from clone RP3-326L13 on chromosome Xq21.1, complete sequence.
ACCESSION Z82170
VERSION Z82170.2 GI:33504458
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Graham, D.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Aug 7, 2003 this sequence version replaced gi:1730463.
Center code: SC
Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP3-326L13 is from the library RPCI-3 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES
source location/Qualifiers
1..127247
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCI326"
/db_xref="taxon:9606"
/chromosome="X"
/map="q21.1"
/clone="RP3-326L13"
/clone_lib="RPCI-3"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 127247;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGATG 20
Db 93812 GTTCATCTCATGTGATG 93793

RESULT 37
LOCUS AP003886 129300 bp DNA linear PLN 21-AUG-2003
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 8, BAC clone:OJ1163_G08, complete sequence.
ACCESSION AP003886
VERSION AP003886.2 GI:34013497
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Euphorbiaceae; Oryzae; Oryza.
REFERENCE
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC clone:OJ1163_G08
JOURNAL Published Only in Database (2001)
AUTHORS 2 (bases 1 to 129300)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kamondai

COMMENT

2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@ias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 On Aug 20, 2003 this sequence version replaced gi:14646819.
 The orientation of the sequence is from M13rev to -21M13 of the BAC
 clone. The nucleotide sequence of this BAC clone was generated by
 combining Monano and RGP-Japan sequencing data.

FEATURES

Location/Qualifiers
 1..129300
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="8"
 /clone="OU1163_G08"

ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 129300;
 Best Local Similarity 90.0%; Pred. No. 1e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
 |||||
 Db 122538 GGTCATCTCATGTTGATG 122557

RESULT 38

AC002089 148416 bp DNA linear PRI 08-OCT-2003
 LOCUS Homo sapiens BAC clone CTA-308B22 from 7, complete sequence.
 AC002089
 AC002089.2 GI:21322223
 HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 148416)
 Sulton, J.E. and Wilson, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 JOURNAL MEDLINE 99063792
 PUBMED 9847074
 2 (bases 1 to 148416)
 REFERENCE
 AUTHORS Hawking, M. and Wollam, A.
 TITLE The sequence of Homo sapiens BAC clone CTA-308B22
 JOURNAL Unpublished (2001)
 3 (bases 1 to 148416)
 REFERENCE
 AUTHORS Waterston, R.
 TITLE Direct Submision
 JOURNAL Submitted (13-MAY-1997) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 4 (bases 1 to 148416)
 REFERENCE
 AUTHORS Waterston, R.
 TITLE Direct Submision
 JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 5 (bases 1 to 148416)
 REFERENCE
 AUTHORS Waterston, R.
 TITLE Direct Submision
 JOURNAL Submitted (04-FEB-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 6 (bases 1 to 148416)
 REFERENCE
 AUTHORS Waterston, R.
 TITLE Direct Submision
 JOURNAL Submitted (04-JUN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 7 (bases 1 to 148416)
 REFERENCE
 AUTHORS Waterston, R.
 TITLE Direct Submision
 JOURNAL Submitted (29-APR-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 8 (bases 1 to 148416)
 REFERENCE

AUTHORS

TITLE

JOURNAL

Wilson, R.
 Direct Submision
 Submitted (08-OCT-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jun 4, 2002 this sequence version replaced gi:2085776.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: saplens@wustl.wustl.edu
 ----- Summary Statistics

 Center project name: H_RG308B22

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NHGRI Chromosome 7 Mapping
 Project and the Washington University Genome Sequencing Center. For
 additional information about the map position of this sequence, see
 http://www.nhgri.nih.gov/DIR/CTB/CHR7 or
 mailto:egreen@nhgri.nih.gov

SOURCE INFORMATION:

Clone CTA-308B22 is from a release of the human BAC library
 CITB-HS-A. The library contains cloned DNA from human sperm. See:
 Shibuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J.
 Kim et al., Genomics 34:213-8 (1996). The clone is available from
 Research Genetics, Inc. (http://www.reagen.com).

VECTOR: pBeltoBAC11

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of CTA-308B22
 actual end is at base position 148416 of CTA-308B22.

FEATURES

source

1..148416
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"
 /clone="CTA-308B22"
 /clone_1lb="CITB-HS-A"
 746..970
 /rpt_family="MIR"
 repeat_region
 1493..2094
 /rpt_family="L1"
 repeat_region
 2095..3501
 /rpt_family="L1"
 repeat_region
 3534..3822
 /rpt_family="Alu"
 repeat_region
 4155..4271
 /rpt_family="Alu"
 repeat_region
 4284..4335
 /rpt_family="MALR"
 repeat_region
 4341..4390
 /rpt_family="MALR"
 repeat_region
 5096..5484
 /rpt_family="L2"
 repeat_region
 5575..6074

```

repeat_region      /rpt_family="L1"
                    6726. .6749
repeat_region      /rpt_family="AT_rich"
                    6981. .7031
repeat_region      /rpt_family="MER103"
                    7661. .7691
repeat_region      /rpt_family="(T)n"
                    8084. .8118
repeat_region      /rpt_family="AT_rich"
                    8449. .8642
repeat_region      /rpt_family="MIR"
                    8659. .8709
repeat_region      /rpt_family="(CAT)n"
                    8717. .9264
repeat_region      /rpt_family="L1"
                    9343. .9406
repeat_region      /rpt_family="AT_rich"
                    9778. .10068
repeat_region      /rpt_family="Alu"
                    10072. .10128
repeat_region      /rpt_family="L2"
                    10190. .10217
repeat_region      /rpt_family="AT_rich"
                    10360. .10560
repeat_region      /rpt_family="L1"
                    11312. .11379
repeat_region      /rpt_family="(TG)n"
                    11335. .12092
repeat_region      /rpt_family="L1"
                    12524. .12548
repeat_region      /rpt_family="AT_rich"
                    13448. .13780
repeat_region      /rpt_family="L1"
                    14263. .14289
repeat_region      /rpt_family="(A)n"
                    15057. .15212
repeat_region      /rpt_family="MIR"
                    15481. .15741
repeat_region      /rpt_family="Alu"
                    15820. .15844
repeat_region      /rpt_family="AT_rich"
                    16142. .16424
repeat_region      /rpt_family="L1"
                    16425. .16817
repeat_region      /rpt_family="L1"
                    16830. .17069
repeat_region      /rpt_family="L1"
                    17070. .17369
repeat_region      /rpt_family="Alu"
                    17370. .17610
repeat_region      /rpt_family="L1"
                    17800. .18064
repeat_region      /rpt_family="L1"
                    18485. .18505
repeat_region      /rpt_family="AT_rich"
                    19665. .19938
repeat_region      /rpt_family="L2"
                    20337. .20661
repeat_region      /rpt_family="MER1_type"
                    20779. .20822
repeat_region      /rpt_family="MALR"
                    20823. .21210
repeat_region      /rpt_family="MALR"
                    21436. .21743
repeat_region      /rpt_family="Alu"
                    21841. .22009
repeat_region      /rpt_family="L1"
                    22014. .22370
repeat_region      /rpt_family="MALR"
                    22609. .22827
repeat_region      /rpt_family="L1"
                    24733. .24754
repeat_region      /rpt_family="AT_rich"

```

```

repeat_region      25752. .25966
                    /rpt_family="L1"
repeat_region      26145. .26179
                    /rpt_family="AT_rich"
repeat_region      26183. .26454
                    /rpt_family="Alu"
repeat_region      26540. .26801
                    /rpt_family="MALR"
repeat_region      26830. .26989

Query Match      84.0%; Score 16.8; DB 9; Length 148416;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGTCCATCTCATGTGTATG 20
        ||||| ||||| |||||
Db      57746 GGTCCATCTCTGTGTATG 57765

RESUT 39
AC011144
LOCUS      Homo sapiens clone RP11-1024, WORKING DRAFT SEQUENCE, 5 unordered
           pieces.
ACCESSION      AC011144.4 GI:13357401
VERSION      HTG, HTGS, PHASE1, HTGS_DRAFT.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 152635)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome, clone RP11-1024
JOURNAL      Unpublished
AUTHORS      2 (bases 1 to 152635)
REFERENCE      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckery,R., Boguslavskiy,L., Boukagalter,B.,
Brown,A., Casale,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galanter,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lhoczeky,J., Lieu,C., Locke,K., Macdonald,P., Margulis,N.,
McSwan,P., McGurk,A., McKernan,K., McLaughlin,J., Weidtm,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 16, 2001 this sequence version replaced gi:12583837.
All repeats were identified using RepeatMasker.
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIPR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2517
Center clone name: 1 O 24
----- Summary Statistics
Sequencing vector: M13: M77815; 3% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 151169 bases at least Q40
Consensus quality: 151789 bases at least Q30
Consensus quality: 152081 bases at least Q20

```

Insert size: 17000; agarose-fp
 Insert size: 152235; sum-of-contigs
 Quality coverage: 6.2 in Q20 bases; agarose-fp
 Quality coverage: 6.9 in Q20 bases.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact size of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 21727: contig of 21727 bp in length
 * 21728 21827: gap of 100 bp
 * 21828 69977: contig of 48150 bp in length
 * 69978 70077: gap of 100 bp
 * 70078 111601: contig of 41524 bp in length
 * 111602 111702: gap of 100 bp
 * 111702 146223: contig of 34522 bp in length
 * 146224 146324: gap of 100 bp
 * 146324 152635: contig of 6312 bp in length.

FEATURES
 source
 1..152635
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-1024"
 /clone_1lb="RP11-1024"
 /clone_1lb="RP11-1024"
 1..21727
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left"
 misc_feature
 21828..69977
 /note="assembly_fragment"
 70078..111601
 /note="assembly_fragment"
 111702..146223
 /note="assembly_fragment"
 146324..152635
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 2; Length 152635;
 Best Local Similarity 90.0%; Pred.No.1e+03; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCATCTCATGTGTGATG 20
 |||||
 Db 7397 GGTCATCTCTTGTGTGATG 7416

RESULT 40
 AC146044/c 155646 bp DNA linear PRI 18-DEC-2003
 LOCUS AC146044
 DEFINITION Pan troglodytes BAC clone RP43-109N23 from 7, complete sequence.
 ACCESSION AC146044
 VERSION AC146044.2 GI:38424257
 KEYWORDS HTG.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 REFERENCE
 AUTHORS 1 (bases 1 to 155646)
 TITLE The sequence of Pan troglodytes BAC clone RP43-109N23
 JOURNAL Unpublished (2001)
 REFERENCE 2 (bases 1 to 155646)
 AUTHORS Wilson,R.
 TITLE Sequencing of Pan troglodytes
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 155646)

AUTHORS
 TITLE Wilson,R.K.
 JOURNAL Direct Submission
 Submitted (01-ANG-2003) Genetics, Genome Sequencing Center, 4444
 Forest Park Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 155646)
 REFERENCE
 AUTHORS Wilson,R.K.
 JOURNAL Direct Submission
 Submitted (19-NOV-2003) Genetics, Genome Sequencing Center, 4444
 Forest Park Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 155646)
 REFERENCE
 AUTHORS Wilson,R.
 TITLE Direct Submission
 Submitted (18-DEC-2003) Department of Genetics, Washington
 University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Nov 19, 2003 this sequence version replaced gi:33386996.
 COMMENT
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics

 Center project name: C_P1109N23

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:
 The RP11-43 BAC library has been constructed by Chung-Li Shu. DNA
 was isolated from white blood cells obtained from a male chimpanzee
 (Pan troglodytes, 'Cint', Yerkes #C0477; birthdate: 6-6-80). The
 clone and detailed information can be obtained from Regen
 (http://www.regen.com) or Pieter de Jong and co-workers at
 http://www.bacpac.chori.org.

NEIGHBORING SEQUENCE INFORMATION:
 This sequence is the entire insert of the clone.

FEATURES
 source
 1..155646
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /chromosome="7"
 /map="7"
 /clone="RP43-109N23"
 /clone_1lb="RP11-43"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 9; Length 155646;
 Best Local Similarity 90.0%; Pred.No.1e+03; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCATCTCATGTGTGATG 20
 |||||
 Db 106829 GGTCATCTCTTGTGTGATG 106810

RESULT 41
 AP005233/c
 LOCUS 157200 bp DNA linear PRI 10-AUG-2002
 DEFINITION Homo sapiens genomic DNA, chromosome 11 clone:RP11-211G23, complete sequence.
 ACCESSION AP005233
 VERSION AP005233.3 GI:22202840
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.
 TITLE Homo sapiens genomic DNA
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 157200)
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (21-MAY-2002) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Saito-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
 COMMENT On Aug 9, 2002 this sequence version replaced gi:21321327.
 FEATURES
 source
 1. 157200
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="RP11-211G23"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 9; Length 157200;
 Best Local Similarity 90.0%; Pred. No. 1e+03; Mismatches 0; Gaps 0;
 Matches 18; Conservative 0; Indels 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
 DB 87060 GGACCATCTCATGTTGATG 87041

RESULT 42
 OSJN00128
 LOCUS 164572 bp DNA linear PLN 30-NOV-2003
 DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0088122, complete sequence.
 ACCESSION AL607001
 VERSION AL607001.4 GI:38567904
 KEYWORDS HTG.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
 AUTHORS Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J., Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, Y., Meng, Q., Zhang, L., Lu, Y., Mu, Y., Lu, Y., Zhang, L., S., Yu, Z., Fan, D., Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J., Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R., Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J., Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q., Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S., Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J., Li, J., Hong, G., Xue, Y. and Han, B.
 TITLE Sequence and analysis of rice chromosome 4
 JOURNAL Nature 420 (6913), 316-320 (2002)
 MEDLINE 22337377

PUBMED 12447439
 REFERENCE
 AUTHORS Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J., Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.U., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K., Zhou, B., Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.O., Guan, J.P., Fu, G., Wang, S.Y., Ren, S.X., Lv, G., Jin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F., Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y., Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J., Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and Hong, G.F.
 TITLE Direct Submission
 JOURNAL Submitted (26-SEP-2001) Han Bin, National Center for Gene Research, Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhan@ncgr.ac.cn
 REMARK bhan@ncgr.ac.cn
 Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSJNBa0088122.
 COMMENT On Nov 30, 2003 this sequence version replaced gi:32492239.
 Web site: http://www.ncgr.ac.cn
 ----- Summary Statistics -----
 Assembly program: phrap

 This is a complete sequence.
 Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (http://www.sottberry.com/), genSCAN (http://CCR-081.mit.edu/GENSCAN.html), GenemarkHMM (http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the complete sequence against NCBI nonredundant protein database (nr) (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR. Location/Qualifiers
 1. 164572
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /varity="Nipponbare"
 /sub_species="japonica"
 /db_xref="taxon:39947"
 /chromosome="4"
 /clone="OSJNBa0088122"
 /clone_1ib="CUGI-OSJNBa"
 /complement (3174..3644)
 /gene="OSJNBa0088122.1"
 /complement (3174..3644)
 /gene="OSJNBa0088122.1"
 /codon_start=1
 /protein_id="CAD1569.2"
 /db_xref="GI:38567905"
 /translation="MAAVAVDVARALRGPDALNFPDILASRTAPPASSADDIRAA
 AAAAAALQHDRAAGGIAIPASGSAHQDQSSAAARTYASGGAQDEGSSGAAGSH
 QYFLDEALFETPOFLRNMAAGMMSPRLSPSSSDSPSPSEAGSGLSMYRDP"
 7871..8788
 /gene="OSJNBa0088122.2"
 /join (7871..8396,8445..8788)
 /gene="OSJNBa0088122.2"
 /codon_start=1
 /protein_id="CAD1570.2"
 /db_xref="GI:38567906"
 /translation="WRBEGEKGVFSFALSGKDEADGKLTDRERRRRRPPDEKD
 KRPGRARAGRGKGRRTLRDLDAALRQPSGAWRTLRGEGAEKMGCGGRTG
 WRKVLGRGGAAGBEATVTLFQHGRTKWDIRGARGLGYRKYVLGWSKAVRRRGR
 REMWDGADALGSGSGGPPRCVSGSASBHKRPDSGGAPLPGREKGVARRAR
 PFRWEGRGHEHCIALDARGVGRARBDADDCCKODCLEASTDTFVLSQAT
 TEGDVSHPRIPE"
 complement (126784..28809)
 /gene="OSJNBa0088122.3"
 /complement (126784..27108,27643..27740,28624..28809)
 /gene="OSJNBa0088122.3"
 /codon_start=1
 /protein_id="CAD1571.2"

gene
CDS
/db_xref="GI:38567907"
/db_xref="SPTREMBL:Q7XU50"
/translation="MLGSLINVSALGFGMTAAVAFCARFVCCARRADASARP
HSPVDFADPFDFDRPIEHSRSGLEPAVAIPTMKNCFASESDTQSCISE
YKEDILRIVPICHNFHLYCLDAWLKOTPCIRISKLPDGKSTVSAFPTMSQP
PILPSSVNPSTHPLPVHDEHSHODGPDMPSEVVEIRQ"
30110..30921
/gene="OSJNB0088122.4"
/join(30110..30283,30363..30683,30769..30921)
/gene="OSJNB0088122.4"
/codon_start=1
/protein_id="CAD41572.3"
/db_xref="GI:38567908"
/db_xref="SPTREMBL:Q7XU49"
/translation="MARRGVDPNWRSDPCVYLTPMCORPLYCGDRCOLASRNP
NIRGRFPFCYRETRTACAYIEWDTEPNVLDITTCLOEGRWYFASSTEQYQK
NAVEROCROSDMRVLTALPMEARPCRCGDCQVLRISNPTLIRFPVYNIL
DDFMBPFRRCYRREIDTRVLTTPPSRVQLELFEQYRVTKARERGEBSRRG"
complement(43323..45005)
/gene="OSJNB0088122.5"
complement(43323..45005)
/gene="OSJNB0088122.5"
/codon_start=1
/protein_id="CAD41573.3"
/db_xref="GI:38567909"
/db_xref="SPTREMBL:Q7XU48"
/translation="MTTCADDTQCAFFAPLLSSKGAENVLLVADDEAEEOQAPVLT
SKPRLAKAWEASVSLGAVPTPMTFTCSARGEARSLGLAFPMILITLILYL
SMISMLFGLGGLAGSLAIGFANITGVSLGLAMGEPICGOAFAGANVALLG
HVRITVLTQGSINLPLTVCGALAIHILPINVLYLVGLGKVALASLVANLNLVL
FLAYIFLKGVHRTGFLLSAESFRGWELSLPSCVSCLEMYEIMLILGL
LNPQATVASMGLIQTSLIYIPSSLSFGSTVSNELGQDEESRATVGLVL
GFGFAPASAPFLVRNWSMFTADPAIVALTAVLPLIGCEHGCNCPQTCGVLR
GSARPKDASINLRSFVLGTVPALVMAWFLDFGLMFLGLAAOACTVYMLLVIG
PTDMAEAKRSQLTGAGANMSDDRAADKSLPVDTPDVERSDHTDRC"
60927..63084
/gene="OSJNB0088122.6"
/join(60927..61265,62629..63084)
/gene="OSJNB0088122.6"
/codon_start=1
/protein_id="CAD41574.3"
/db_xref="GI:38567910"
/db_xref="SPTREMBL:Q7XU47"
/translation="MDEIGPSPRGEAEAGDLTSTAREAGROGAGAVRASQAPSV
YOOPATNGMGACVPGPHIIVFORJSHKALIALPCIIAANKAFVHTSSYKTR
LILISTESBOLQLPPODRORCDROLISYRALMEMWMLCGTTAVGPPPEPLSS
SSSINQSRFPIDSPYQAVTSTYLMERPSPRSRDSSTSYVANKSDASTOREVMM
LQGRRTVTTVTRKMGPTMAPLITTSKTYISTNFRGIKVDIVL"
complement(63664..66067)
/gene="OSJNB0088122.7"
complement(63664..66067)
/gene="OSJNB0088122.7"
/codon_start=1
/protein_id="CAD41575.3"
/db_xref="GI:38567911"
/db_xref="SPTREMBL:Q7XU46"
/translation="MPESCVAALVLYLDPEICKARLNRARFGAASADCVAAKLPA
NRYRALAAADDSDGCGATENGSRSSAAMIKEEYALICPTPEDGKTFEPM
EKNGKICISISKMATIGIDRRYKWSLSTESRPHNVALLOOTIMEVAGEIDFC
EPAGSYSLFRLQGLGPHKRYGRVYGHSHGMNKPRFPLSTSDDOATSQYYLN
EPGNMILYHVGDFVSSDQLTNLKFSMWQIDCTHKGGLCVDVSFIYPKGRHBDCT
ICK"
complement(70237..74263)
/gene="OSJNB0088122.8"
complement(70237..74263)
/gene="OSJNB0088122.8"
/codon_start=1
/protein_id="CAD41576.3"
/db_xref="GI:38567912"
/db_xref="SPTREMBL:Q7XU45"
/translation="MKEKTLAVSGLPSARMRQSRBASWYQLSPMTCVACTSTLVS

gene
CDS
/db_xref="GI:38567913"
/db_xref="SPTREMBL:Q7XU44"
/translation="MVAASPPHSSSLFSLHLRTSTVARATGMRGREGSRROY
MRRRRSASRARPSPSPGRSMATLAALSLPGAALAAEAERRRVYSGSQRLARGA
YDNPPTKODSNISHVENFGCVLCRAMKTNLANCASXPYSEQFPGTPY"
80941..86909
/gene="OSJNB0088122.10"
/join(80941..82090,82295..82536,82597..84075,84346..84484,
86395..86640,86692..86909)
/gene="OSJNB0088122.10"
/codon_start=1
/protein_id="CAD41578.3"
/db_xref="GI:38567914"
/db_xref="SPTREMBL:Q7XU43"
/translation="WRLEKAVESWIPAVEKKVDVLTVDLQVNOQSSAMKT
EQTPNAGLDPGRVNGSTGLDGRVQILNREBAGFTPEKIPANGTYASHVNSPLP
FHLDAASNSFOSLVSQMGVAPMTCPEKEDGNPTMWSNCEAYPDVGIHGNWY
KVATLNFGNANAFILQAVRSQLEGVTSSELCIKTRFNKROQALIRWIKICQTS
MEWHDTTKKEVYVCTMNAFLPMIENYGGCFEDDVDAQLVDOEYVLOLGN
YGTSSKTHSNPSSSYSHGELGRRKSPASVAYQLVDEALRELOQMDULASAS
IDHNILHEHGKPLASSISGSGNASARPPVYVNDGIDPNNMYEELQDGAIGTE
SKGPEIDVILPSTYKIRIFSRKEKHDECVICMWTYKONDRLTKLPCRHQYHOTCV
TKMKIKINVCVCKEYVSGSK"
75062..77788
/gene="OSJNB0088122.9"
/join(75062..75363,76561..76657,77720..77788)
/gene="OSJNB0088122.9"
/codon_start=1
/protein_id="CAD41577.3"
/db_xref="GI:38567913"
/db_xref="SPTREMBL:Q7XU44"
/translation="MSAASPPHSSSLFSLHLRTSTVARATGMRGREGSRROY
MRRRRSASRARPSPSPGRSMATLAALSLPGAALAAEAERRRVYSGSQRLARGA
YDNPPTKODSNISHVENFGCVLCRAMKTNLANCASXPYSEQFPGTPY"
80941..86909
/gene="OSJNB0088122.10"
/join(80941..82090,82295..82536,82597..84075,84346..84484,
86395..86640,86692..86909)
/gene="OSJNB0088122.10"
/codon_start=1
/protein_id="CAD41578.3"
/db_xref="GI:38567914"
/db_xref="SPTREMBL:Q7XU43"
/translation="WRLEKAVESWIPAVEKKVDVLTVDLQVNOQSSAMKT
EQTPNAGLDPGRVNGSTGLDGRVQILNREBAGFTPEKIPANGTYASHVNSPLP
FHLDAASNSFOSLVSQMGVAPMTCPEKEDGNPTMWSNCEAYPDVGIHGNWY
KVATLNFGNANAFILQAVRSQLEGVTSSELCIKTRFNKROQALIRWIKICQTS

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 8; Length 164572;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20
Db 113198 GGTACATACGATGTTGATG 113217

RESULT 43
SPNEU908/c
LOCUS SPNEU908 175936 bp DNA linear HTG 11-JUL-2001
DEFINITION Streptococcus pneumoniae clone G54, *** SEQUENCING IN PROGRESS ***.
ACCESSION AL449930
VERSION AL449930.1 GI:11545155
KEYWORDS HTG; HTGS PHASE2.
SOURCE
ORGANISM Streptococcus pneumoniae
Bacteria; Filimutices; Lactobacillales; Streptococcaceae;
Streptococcus.
1
Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Humbert,Y.,
Friedl,L., Guerrier,M., Grand-Schenk,B., Gandin,C., de
Francesco,M., Polissi,A., Buell,G., Feger,G., Garcia,B., Peltech,M.
and Garcia-Buñtos,J.F.
Annotated draft genomic sequence from a Streptococcus pneumoniae
type 19F clinical isolate
Microb. Drug Resist. 7 (2), 99-125 (2001)
2135329
11442348
2 (bases 1 to 175936)
Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A.,
Humbert,Y., Friedl,L., Guerrier,M., Grand-Schenk,B., Gandin,C., de
Francesco,M., Buell,G., Feger,G., Garcia,E., Peltech,M. and
Garcia-Buñtos,J.F.
Direct Submission
Submitted (31-OCT-2000) Research Department, Glaxo Wellcome, S.A.,
Severo Ochoa 2, 28760 Tres Cantos, SPAIN
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

* provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 175936: contig of 175936 bp in length.
 Location/Qualifiers
 1. 175936
 /organism="Streptococcus pneumoniae"
 /mol_type="genomic DNA"
 /serotype="19F"
 /db_xref="taxon:1313"
 /clone="G54"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 175936;
 Best Local Similarity 90.0%; Pred. No. 1.1e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCGTACTACTGATGATG 20
 |||||
 Db 29955 GGTCCATCTCTGATGATG 29936

RESULT 44
 AC139894/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-278P24, *** SEQUENCING IN PROGRESS
 AC139894
 AC139894.1 GI:28394806
 HTG: HTGS PHASE1.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 186184)
 Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Barnmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guertiz, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kows, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensheva, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,
 Matheswaril, M., Mahindaratne, M., Mamoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawlin, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
 Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
 Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
 Newson, N., Nguyen, N., Norris, S., Nwackel, O., Okunolu, G.,
 Olarinmusa, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
 Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,
 Popovic, D., Primus, E., Pu, L., Plazo, M., Quiroz, J., Raculin, E.,
 Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
 Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

Project Information
 Center project name: KOND
 Center clone name: CH230-278P24
 Sequencing Vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 140885 bases at least Q40
 Consensus quality: 149547 bases at least Q30
 Consensus quality: 158103 bases at least Q20
 Estimated insert size: 156705; sum-of-contigs estimation
 Quality coverage: 2x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 70 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 1154: contig of 1154 bp in length
 1155 1254: gap of unknown length
 1155 1283: contig of 1029 bp in length
 1284 2383: gap of unknown length
 2384 3446: contig of 1063 bp in length
 3447 3546: gap of unknown length
 3547 5270: contig of 1723 bp in length
 5270 5369: gap of unknown length
 5370 6862: contig of 1493 bp in length
 6863 6962: gap of unknown length
 6963 8194: contig of 1232 bp in length
 8194 8294: gap of unknown length
 8295 10111: contig of 1817 bp in length
 10111 10211: gap of unknown length
 10212 11586: contig of 1375 bp in length
 11587 11686: gap of unknown length
 11687 13117: contig of 1431 bp in length
 13117 13218: gap of unknown length
 13218 14697: contig of 1480 bp in length

Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Savery, G., Scherer, S.,
 Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A.,
 Sisson, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E.,
 Song, X.-Z., Sorrell, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
 Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
 Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D.,
 Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
 Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R.,
 Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
 Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
 Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,
 Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 186184)
 Worley, K. C.
 Direct Submission
 Submitted (16-FEB-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 186184)
 Worley, K. C.
 Direct Submission
 Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

```

* 14698 14797: gap of unknown length
* 14798 16013: contig of 1216 bp in length
* 16014 16113: gap of unknown length
* 16114 17514: contig of 1401 bp in length
* 17515 17614: gap of unknown length
* 17615 18912: contig of 1298 bp in length
* 18913 19012: gap of unknown length
* 19013 20041: contig of 1029 bp in length
* 20042 20141: gap of unknown length
* 20142 21580: contig of 1439 bp in length
* 21581 21680: gap of unknown length
* 21681 23564: contig of 1884 bp in length
* 23565 23664: gap of unknown length
* 23665 24769: contig of 1105 bp in length
* 24770 24869: gap of unknown length
* 24870 26701: contig of 1832 bp in length
* 26702 26801: gap of unknown length
* 26802 28355: contig of 1554 bp in length
* 28356 28455: gap of unknown length
* 28456 29847: contig of 1392 bp in length
* 29848 29947: gap of unknown length
* 29948 31194: contig of 1247 bp in length
* 31195 31294: gap of unknown length
* 31295 32553: contig of 1259 bp in length
* 32554 32653: gap of unknown length
* 32654 34999: contig of 2346 bp in length
* 35000 35099: gap of unknown length
* 35100 36223: contig of 1124 bp in length
* 36224 36323: gap of unknown length
* 36324 38111: contig of 1788 bp in length
* 38112 38211: gap of unknown length
* 38212 40164: contig of 1953 bp in length
* 40165 40264: gap of unknown length
* 40265 42432: contig of 2168 bp in length
* 42433 42532: gap of unknown length
* 42533 44066: contig of 1534 bp in length
* 44067 44166: gap of unknown length
* 44167 46233: contig of 2067 bp in length
* 46234 46333: gap of unknown length
* 46334 48687: contig of 2254 bp in length
* 48688 48887: gap of unknown length
* 48888 50591: contig of 1904 bp in length
* 50592 50691: gap of unknown length
* 50692 52837: contig of 2146 bp in length
* 52838 52937: gap of unknown length
* 52939 55123: contig of 2186 bp in length
* 55124 55223: gap of unknown length
* 55224 56867: contig of 1644 bp in length
* 56868 56967: gap of unknown length
* 56968 58897: contig of 1930 bp in length
* 58898 58997: gap of unknown length
* 58998 61198: contig of 2201 bp in length
* 61199 61298: gap of unknown length
* 61299 62820: contig of 1522 bp in length
* 62821 62920: gap of unknown length
* 62921 64789: contig of 1869 bp in length
* 64790 64889: gap of unknown length
* 64890 66789: contig of 1900 bp in length
* 66790 66889: gap of unknown length
* 66890 69075: contig of 2186 bp in length
* 69076 69175: gap of unknown length
* 69176 70324: contig of 1149 bp in length
* 70325 70424: gap of unknown length
* 70425 73643: contig of 3219 bp in length
* 73644 73743: gap of unknown length
* 73744 75986: contig of 2243 bp in length
* 75987 76086: gap of unknown length
* 76087 78980: contig of 2894 bp in length
* 78981 79080: gap of unknown length
* 79081 81809: contig of 2723 bp in length
* 81810 84998: contig of 3088 bp in length
* 84999 85098: gap of unknown length

```

```

Query Match
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGTCCATCTCATGTTGATG 20
Db 169681 GCCCTACTCATGTTGATG 169662

RESULT 45
AC119611/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULFILL.
Rattus norvegicus (Norway rat)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 209706)
Muzny,D,Marie, Metcalf,M, Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alebrooks,S, Amin,A, Anguiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Ceaar,H, Center,Z,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denson,S, Dexamio,C, Ding,Y, Dinh,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Georgievski,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
Guevarre,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M,
Hollins,B, Howells,S, Huylk,S, Hume,J, Idlebird,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,
Kowis,C, Kraft,C,L, Lebow,H, Levan,Z, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
Lorenshewa,L, Louised,H, Lorado,R,S, Lu,X, Ma,J,
Maheshwari,M, Mahindartine,M, Mahmood,M, Malloy,K, Mangum,A,
Mangum,B, Mapa,P, Martin,K, Martin,R, Martinez,E,
Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
Nwankweliem,O, Okwoum,G, Olarpunagoon,A, Pal,S, Parks,K,
Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,L,
Plapper,F, Poindexter,A, Popovic,D, Primus,B, Pu,L,
Punzo,M, Quiroz,J, Rachlin,E, Reeves,K, Reiter,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
Rivers,C, Rodery,T, Rojce,A, Rose,M, Rose,R, Ruiz,S,J,
Sander,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H,
Shetty,J, Shwartsbeyn,A, Sisson,I, Sitter,C,D, Snares,D,
Sneed,A, Sodergren,E, Song,X-Z, Sorille,R, Sosa,J,
Steinle,M, Strong,R, Sutton,A, Swalek,A, Taber,P, Taylor,C,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K,
Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J,
Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,

```


Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 209706)
Worley, K.C.

Direct Submission
Submitted (29-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 209706)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23808767.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYON
Center clone name: CH230-440C15

----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 187370 bases at least Q40
Consensus quality: 188624 bases at least Q30
Consensus quality: 189405 bases at least Q20
Estimated insert size: 192639; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

- * NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
- * This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
- * 1 209706: contig of 209706 bp in length.

FEATURES

source
1. 209706
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-440C15"
2286. 3117
/note="clone boundary
clone_end:5p6
site:
end_sequence:B2128438"
202827. 203690
/note="clone boundary
misc_feature
misc_feature
misc_feature

clone_end:77
site:
end_sequence:B2128438"
206213. 207242
/note="wgs_end_extension
clone_end:77"
207915. 209706
/note="wgs_end_extension
clone_end:77"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 209706;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
Db 98490 GGTCCATCTCATGTTGATG 98471

RESULT 46

AC129598/c

LOCUS

DEFINITION

Mus musculus BAC clone RP23-28305 from chromosome 18, complete sequence.

AC129598

VERSION

AC129598.4 GI:33285248

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.

1 (bases 1 to 213381)

Isak, A. and Haglund, K.

The sequence of Mus musculus BAC clone RP23-28305

Unpublished (2001)

2 (bases 1 to 213381)

Wilson, R.

Sequencing of Mus musculus

Unpublished (2001)

3 (bases 1 to 213381)

McPherson, J.D. and Waterston, R.H.

Submitted (30-JUL-2002) Genome Sequencing Center, 4444 Forrest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 213381)

Wilson, R.

Submitted (20-MAY-2003) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 213381)

Wilson, R.

Submitted (25-JUL-2003) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 213381)

Wilson, R.

Submitted (25-NOV-2003) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Jul 26, 2003 this sequence version replaced gi:30911163.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: M_BA0283005

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap

between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The Rpci-23 BAC Library has been constructed by Kazutoyo Oseegawa
and Minako Tateo in the laboratory of Pieter de Jong
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

Location/Qualifiers

FEATURES

source

1..213381

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="18"

/map="18"

/clone="RP23-28305"

/clone_lib="RPci-23"

61..124

/rpt_family="MER1_type"

236..338

/rpt_family="L1"

3234..3467

/rpt_family="L1"

4674..4747

/rpt_family="Alu"

5793..5930

/rpt_family="MER1_type"

6211..6610

/rpt_family="L1"

6804..7142

/rpt_family="L1"

9876..10022

/rpt_family="Alu"

11296..11350

/rpt_family="ERV1"

11818..12016

/rpt_family="L1"

12019..13959

/rpt_family="L1"

13961..14028

/rpt_family="L1"

14028..14346

/rpt_family="L1"

14362..14699

/rpt_family="MALR"

14704..14809

/rpt_family="L1"

14832..15990

/rpt_family="L1"

15989..16163

/rpt_family="L1"

16170..16379

/rpt_family="L1"

16380..16811

/rpt_family="MALR"

repeat_region

16831..17134

repeat_region

17112..17202

repeat_region

17195..17789

unsure

19446..19613

repeat_region

24760..24858

repeat_region

24882..26242

repeat_region

26273..27217

repeat_region

27233..27300

repeat_region

27863..27911

repeat_region

28816..29068

repeat_region

31173..31873

repeat_region

32908..34503

repeat_region

38155..39637

repeat_region

42874..42931

repeat_region

42985..43199

repeat_region

43419..43722

repeat_region

43742..44280

repeat_region

45706..45906

repeat_region

46062..46394

repeat_region

46258..46627

repeat_region

48421..48648

misc_feature

48421..48648

repeat_region

48480..48877

repeat_region

49189..49637

repeat_region

49955..50433

repeat_region

50852..50962

repeat_region

50965..51177

repeat_region

51200..52281

repeat_region

52304..52442

repeat_region

52838..52889

repeat_region

52934..53198

repeat_region

53908..54059

repeat_region

56035..56662

repeat_region

56230..56337

unsure

56660..57273

repeat_region

57429..57577

repeat_region

57628..58265

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

repeat_region

59055..59425

```

repeat_region      /rpt_family="L1"
                   59626..60201
                   /rpt_family="L1"

Query Match      84.0%; Score 16.8; DB 10; Length 213381;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGTCCACTCATGTTGATG 20
    |||||
Db 104478 GGTCAACACTCATGTTGATG 104459

RESULT 47
AC107756 216911 bp DNA linear ROD 30-SEP-2003
LOCUS Mus musculus chromosome 18, clone RP23-214K19, complete sequence.
DEFINITION AC107756
ACCESSION AC107756.9 GI:33414530
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 216911)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
JOURNAL Mus musculus chromosome 18, clone RP23-214K19
TITLE Unpublished
AUTHORS 2 (bases 1 to 216911)
          Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
          Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhgalter, B.,
          Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
          Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,
          Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S.,
          Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
          Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
          Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
          Kamat, A., Karatas, A., Kelle, C., LaRoque, K., Lamazares, R.,
          Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
          Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
          McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T.,
          Mlegana, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
          Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
          Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
          Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
          Roerich, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S.,
          Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
          Straus, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
          Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
          Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
          Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 216911)
          Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
          Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
          Boguslavsky, L., Bouckhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
          Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K.,
          Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
          Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
          Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
          Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
          Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
          Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
          Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
          Meldrim, J., Menus, L., Mihova, T., Mlegana, V., Murphy, T., Naylor, J.,
          Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
          O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
          Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
          Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
          Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
          Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
          Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
          Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

```

```

TITLE Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
JOURNAL Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (25-JUN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 216911)
          Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
          Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
          Boguslavsky, L., Bouckhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
          Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K.,
          Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
          Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
          Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
          Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
          Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
          Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
          Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
          Meldrim, J., Menus, L., Mihova, T., Mlegana, V., Murphy, T., Naylor, J.,
          Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
          O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
          Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
          Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
          Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
          Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
          Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
          Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (03-AUG-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 216911)
          Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
          Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
          Boguslavsky, L., Bouckhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
          Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K.,
          Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
          Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
          Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
          Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
          Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
          Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., Maclean, C.,
          Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
          Meldrim, J., Menus, L., Mihova, T., Mlegana, V., Murphy, T., Naylor, J.,
          Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
          O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
          Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
          Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
          Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
          Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
          Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
          Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (30-SEP-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 3, 2003 this sequence version replaced gi:32189542.
All repeats were identified using RepeatMasker:
  smit, A.F.A. & Green, P. (1996-1997)
  http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20399
Center clone name: 214_K_19
-----
Location/Qualifiers
1..216911
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/map="18"

```

FEATURES

SOURCE

```

misc_feature
/clone="RP23-214K19"
/clone_lib="RPCI-23 Female Mouse BAC"
1..2633
/note="wgs end extension
clone_end:SP6"
repeat_region
/rpt_family="L1"
1..783
800..2024
/rpt_family="L1"
repeat_region
2034..3892
/rpt_family="L1_MM"
2634..2639
/note="clone_boundary
clone_end:SP6
site:ECORI"
repeat_region
complement(3908..4155)
/rpt_family="L1"
4156..4196
/rpt_family="7A)n"
repeat_region
complement(4197..4639)
/rpt_family="L1"
repeat_region
complement(4657..5359)
/rpt_family="LX5"
repeat_region
complement(5372..5778)
/rpt_family="L1"
5826..5890
/rpt_family="TATATG)n"
repeat_region
complement(6783..6901)
/rpt_family="MIR3"
repeat_region
complement(7147..7655)
/rpt_family="L1_MM"
repeat_region
complement(7852..7962)
/rpt_family="PB1D10"
8947..9105
/note="single clone coverage"
repeat_region
complement(9226..9405)
/rpt_family="MIR"
repeat_region
complement(9460..9819)
/rpt_family="RLRERN MM"
repeat_region
complement(9820..9976)
/rpt_family="MERSA"
12506..12585
/rpt_family="(TC)n"
repeat_region
complement(13129..13294)
/rpt_family="Lx2"
13295..14013
/rpt_family="RLTR13C"
repeat_region
14017..14180
/rpt_family="RLTR13D"
repeat_region
complement(14181..14269)
/rpt_family="Lx"
14618..14756
/rpt_family="B3"
repeat_region
complement(14856..15358)
/rpt_family="L1"
16496..16848
/rpt_family="RLR10C"
16949..17012
/rpt_family="(TTTC)n"
repeat_region
complement(17015..17261)
/rpt_family="L1_MM"
17480..17686
/rpt_family="Lx8"
17921..19038
/rpt_family="Lx7"
repeat_region
complement(19022..19450)
/rpt_family="L1"
19451..20633
/rpt_family="L1_MM"

```

```

Query Match      84.0%; Score 16.8; DB 10; Length 216911;
Best Local Similarity 90.0%; Pired, No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1  GGTCCATCTCATGTTGATG 20
      |||||
Db      195618  GGTCAACATCATGTTGATG 195637

```

RESULT 48

```

AC107742      225182 bp  DNA  linear  HTG 07-JUN-2003
LOCUS      AC107742
DEFINITION  Mus musculus clone RP23-331D7, WORKING DRAFT SEQUENCE, 9 ordered
pieces.

```

```

AC107742      GI:31455675
VERSION      AC107742.3

```

```

HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

```

Submitted (07-JUN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 7, 2003 this sequence version replaced gi:20303769.

```


SNGERKIENSDDI LAMYLEDAHVACVGSTSGAPECIRMSVATSDENIVEAIRI KEA
LAKLX"
gene complement (1743. . 2957)
/locus_tag="BT2416"
CDS complement (1743. . 2957)
/locus_tag="BT2416"
/codon_start=1
/transl_table=11
/product="GMP cyclohydrolase II"
/protein_id="AAO77523.1"
/db_xref="GI:29339728"
/translation="MBEIMDRIRIEDAIDFKKSGKYVYVDDRENEGDLIIAAKTI
PEKVPFMLKHAGVLCAPITVSRCKELDIPHQVSDNTSVLGPFTVTIDKLEGCTGV
SASDRAATIOALADPASTPATPGRGHINPLAOKRGVLRAGHTEATIDMARLGLY
PAGLAMEITSDDGTADARPELIRAMADEHGLKLSIHDILAYRLKQESIVEKGVENMP
TEHGMERLIPROKSNGLNLEHMAIPKGTWNEDEPIIVRVHSSCATGDIIGSHRCDGEO
LHKAMEITKEKSGKYVYVYINQSGRIGLMEKKKAYKLOEDGMDITYDANUCIHLADER
DHGAQOILREIVGHKMKLLTNNPVPRVGLAAYGLEIVENVEPTEVPDPYNERVLTATK
KERMGHTLHFNK"
gene complement (2963. . 4813)
/locus_tag="BT2417"
CDS complement (2963. . 4813)
/locus_tag="BT2417"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO77524.1"
/db_xref="GI:29339729"
/translation="MLPTGTFPICPIEMMOPIRMYVDELNGKGLDMSVIAOFFEYSA
LTLVPSLPALILASLITFGNGERYELAKKAAGISILKIMPLAPVCGLVAVSF
YQNVNVPQAOTKLYTLIIISMOKSPELDIPGVVSELPIDINLIKAKNRKTGMWYD
VLIVNLDGFENAHIIYADSGHMENTADQKHLMLHYSGLDFENLRSSQMSQNVPR
REEFEKTIIEFDSDFNMAOASIMNSTSAADOMGLQASIDSMKMGVPSIGROYRE
VSEGNRSTYGLTKEDDYKIMEANITENVIDLEVAVALPOMOKOVISSAASAEMLSS
DINRPSYMGSHDYSMRKHEITEMKKTISLISCLIPRTIGALGIIIRKGIIGMVIY
SVWFTIYITINDTYKAKRDRVYVMWGWTSSAYLAPLAFPLYTKSKDSVYUNAD
AVINWPKIIVGIRNIRHLPEKKEVYIHDPTYTIPDLOVLVDECKEYIAKNKMLAPN
YFKLMGSGEKDEMISSINEKLETLIEEMNSTSVLLNTLNNYPIISVSAHTRPHY
WNLAAGVIPFGILFFPYFRIMAFVRALAKDMERIIKNNEQIPIIONINK"
gene complement (4938. . 5324)
/locus_tag="BT2418"
CDS complement (4938. . 5324)
/locus_tag="BT2418"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO77525.1"
/db_xref="GI:29339730"
/translation="MKKEKIHLEYLLNATSKNII LMAAISTPTGLEDMFADKVVSDKI
VEFHMGKTEQRNABEIIAIRSFSFIRFQWDDNERDYFEIKMTYNELTSDVYLEITDF
AEADRVADMKELMESQVAKLRTCGP"
crRNA 5549. . 5622
/product="crRNA-mec"
/note="codon recognized: AUG"
gene complement (5789. . 7096)
/locus_tag="BT2419"
CDS complement (5789. . 7096)
/locus_tag="BT2419"
/codon_start=1
/transl_table=11
/product="putative GMP synthase [glutamine-hydrolyzing]"
/protein_id="AAO77526.1"
/db_xref="GI:29339731"
/translation="MKQDMIVILDGSHENTVLARALAGVYSIYPHDITVEELKA
LENVKGITINGGLANNVIDGAILDVNPSITWTGCIPIVMAAGHDKATCAVUKPATPTDIEA
ITAAIKSVPTDTCQAEANNNMANFVNDQIELIRROYGDKVLLALSGGDSVVAAAL
LKAIGENLVCVHNGLMKRGSSEDEVFVSQULKANLYVYVTDTRFDKLAGVDP
QRAKIGGFIRFVEBEARKLDIDPLGGGIIYPIDIVESGITAVKVVSHNVGGLPE
DLKPOLVEPLROLKDEVRACGLELGPYEMVYRQFPFGGLGVRCLGAIITDRLEAV
RSDADILREEPFAGLDKKVMQYFTVVPDFKSVGVARNRSPDWPIIRAVNTVDMT
ATLEPVDWIIIMKIDRILIKVYKANNRVCIYDMSPKRNATIEWE"
gene complement (7282. . 8730)
/locus_tag="BT2420"

CDS complement (7282. . 8730)
/locus_tag="BT2420"
/codon_start=1
/transl_table=11
/product="two-component system histidine kinase, PrtD"
/protein_id="AAO77527.1"
/db_xref="GI:29339732"
/translation="MNDMNIKTKLILIGLMAGNIILLVTLSSVNLQLLATPEPSPA
AMPALERALLWISVTGGICITLGLTLVWLPRTISKPIKEIKQGIIEIANHNEERLD
MRNOEPREVEVESFNRAERLAEYRASTLADILSAKKEIEAIVSIDDPILGLMDRE
ILPTNEALNIVLNRGNVIRKSAEELSLKNDILRLIRLELVTPGDDQKEPLKIYADK
ESYKASQIPIINTDAGKDEBSKLDVITILKNITERPRELDSAKTPTSTISHEKTPY
AAIMWSQILEDRKVGELNDEQOLSKSIKENSERLSTIGELNMTQVEKGLQMP
KITPPIELIEVALIANOVQADKFNIOIEVEYPERKIGKLFPVDSKRIAMVLTNLSNAI
RYSKENGIVVGAIGQEMDILVYQDFGKGLDPYRHSIFIDRYFRVGTQVQSGGLGL
SISDPYEAHQGTITVESSELGKSGRPFWRKLKA"
gene complement (8741. . 9874)
/locus_tag="BT2421"
CDS complement (8741. . 9874)
/locus_tag="BT2421"
/codon_start=1
/transl_table=11
/product="PrnC, osmosensitive K+ channel histidine kinase"
/protein_id="AAO77528.1"
/db_xref="GI:29339733"
/translation="MDMDREQSVQHFLLDIKKSHRGRKTIYIGMIAVGKSYRMLQEA
HELLENGVDVQIGYIETHGRAGTALMOGLPIPRRKIPYKSKELSEMDVALIRLP
EIVIVDELATHNVEGSLNEKRWQDMVTLLDGIINVAISAINQIHESVNEEVOITGLE
VKEVVPDSVIOEADENVNIDLTABELISRLKAGIYPERKIQTDALPFPRENTLOLR
ELALKENALREKKEVBNVWGVSVGIRHEKFPACISSEKTPRIRIKARLARTRN
TTFVALYQTPRESMDRIIDLASQRYLNLHFLVLELGEVYIQVSKOILGSIYKVCRD
KOISVVCQGTNLEPRPAHICSLQRYRFLNNLSQANVDLIILA"
gene complement (9667. . 10713)
/locus_tag="BT2422"
CDS complement (9667. . 10713)
/locus_tag="BT2422"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO77529.1"
/db_xref="GI:29339734"
/translation="MKSFPFSKTIYVGLVAATVVSQGITDAKQAEFVQGDIVSSYV
RGMYQGGGAAPQPTLGFQDNFSVYANGSTNFSQGNKELDLTLAYKFGKQPTLVAD
LWMBEGAYKYFENFKSHETGHHFEAGLAYTLPEKFPPLSVAMVTMFGKCKKLDSSB
LKOVYSSYLELNPFPESVANNVDLTCGAVPYKABGIYVNSGFAVTNALKGMTIKIN
DKFSLPIFQAQIWNPCVEDAEIVFGVTLRP"
gene complement (10829. . 11398)
/locus_tag="BT2423"
CDS complement (10829. . 11398)
/locus_tag="BT2423"
/codon_start=1
/transl_table=11
/product="K+-transporting ATPase C chain"
/protein_id="AAO77530.1"
/db_xref="GI:29339735"
/translation="MKTLEFKSLITLVCFVFSFYIILMLFAQVAGPNKGAENVAT
LDKRVGAANYGMFTKOIYFMGSPSCAGDGVSSSSGSGKPTNEPYLAEGARID
TPIVHNPYLSKQVPAEMWTASAGLDPNITPQCAVYQVQVRAQARGITENQVEIYD
QSEKPLIGITGTEKINVLKUNILEBK"
gene complement (11415. . 13448)
/locus_tag="BT2424"
CDS complement (11415. . 13448)
/locus_tag="BT2424"
/codon_start=1
/transl_table=11
/product="K+-transporting ATPase B chain"
/protein_id="AAO77531.1"
/db_xref="GI:29339736"
/translation="MKMKKSASLPQKGOVIESLKQSFVKLNPRMLIKNPIYTVEVAT
VVMFLVTLYSTIINSQGSFAYNIAVPIILPTLLFAPPAEIIAARGAQAQDSLRKTR
ETPRAKKVEGNQIVTSSQLKGDVVFCEAGDVIPSDGEIIEGLASIDESAITGESA

Query Match 84.0%; Score 16.8; DB 1; Length 300340;
 Best Local Similarity 90.0%; Pred. No. 1.1e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTGATG 20
 |||||
 DB 114904 GGTCCTACTCATGTGATG 114885

RESULT 50

AX571760/c AX571760 349980 bp DNA linear PAT 29-NOV-2002
 LOCUS AX571760
 DEFINITION Sequence 4979 from Patent WO02077021.
 ACCESSION AX571760
 VERSION AX571760.1 GI:26003952
 KEYWORDS

SOURCE Streptococcus pneumoniae
 ORGANISM Streptococcus pneumoniae
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.

REFERENCE

1 Masignani,V., Tetrelin,H. and Fraser,C.
 AUTHORS Streptococcus pneumoniae proteins and nucleic acids
 TITLE Patent: WO 02077021-A 4979 03-OCT-2002;
 JOURNAL Chiron Spa (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
 Chiron Spa (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
 Location/Qualifiers

FEATURES

source 1.
 1. 349980
 /organism="Streptococcus pneumoniae"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1313"
 /note="seq 4979 too long: 2.162.598 bases-replaced by
 following seq:-seq 4979: from 0.000.001 to 0.349.980-seq
 4980: from 0.300.001 to 0.649.980-seq 4981: from 0.600.001
 to 0.949.980-seq 4982: from 0.900.001 to 1.249.980-seq
 4983: from 1.200.001 to 1.549.980-seq 4984: from 1.500.001
 to 1.849.980-seq 4985: from 1.800.001 to 2.149.980-seq
 4986: from 2.100.001 to 2.162.598"

ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 349980;
 Best Local Similarity 90.0%; Pred. No. 1.1e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTGATG 20
 |||||
 DB 67573 GGTCCTACTCATGTGATG 67554

Search completed: August 17, 2004, 15:02:53
 Job time : 1612 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 13:18:52 ; Search time 360 Seconds
(without alignments)
236.011 Million cell updates/sec

Title: US-09-825-489-3
Perfect score: 20
Sequence: 1 ggtccatctactgtgatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 800 summaries

Database : N_Geneseq_29aug04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002s:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	4	AAD21384 Antisense
2	20	100.0	486	6	ABL84566 Human ova
3	20	100.0	506	6	ABL79139 Human ova
4	20	100.0	513	6	ABL79129 Human ova
5	20	100.0	822	4	AAS63227 Human xer
6	20	100.0	1377	4	AAD21394 Human xer
7	20	100.0	1407	4	AAS34872 CDNA enco
8	20	100.0	1407	6	ADC46030 Human neo
9	20	100.0	4670	6	ABZ35476 Human gen
10	16.8	84.0	1029	7	ABX05799 S. pneumo
11	16.8	84.0	14273	2	AAVS2173 AAVS2173
12	16.8	84.0	110000	7	ABSS56454_00
13	16.4	82.0	6765	7	ABT13966 Human Ctr
14	16.4	82.0	15246	7	ABT13954 Human Ctr
15	16.4	82.0	110000	6	ABA90521_15
16	15.8	79.0	431	6	ABV97710 Human pan
17	15.8	79.0	462	8	ACH25004 Human adu
18	15.8	79.0	614	3	AAA16057 Human col
19	15.8	79.0	2830	4	ABL23928 Drosophil
20	15.8	79.0	6454	4	ABL11966 Drosophil
21	15.8	79.0	8114	4	ABL03787 Drosophil
22	15.8	79.0	18737	4	ABL03786 Drosophil
23	15.8	79.0	24401	4	ABL03396 Drosophil

24	15.4	77.0	551	4	ABK42099 CDNA enco
25	15.4	77.0	551	8	ADB59766 Connectiv
26	15.4	77.0	609	5	ABV55825 Human pro
27	15.4	77.0	633	4	AAH52789 S. epider
28	15.4	77.0	675	6	AAH53645 S. epider
29	15.4	77.0	675	6	ABN90925 Staphyloc
30	15.4	77.0	2132	3	AAC46198 Arabidops
31	15.4	77.0	2135	3	AAC40444 Arabidops
32	15.4	77.0	2946	4	AAH54374 S. epider
33	15.4	77.0	3021	4	AAH54380 S. epider
34	15.4	77.0	3391	4	AAH54393 S. epider
35	15.4	77.0	7722	5	AAH85546 DNA enco
36	15.4	77.0	28564	9	ADD46613 Human gen
37	15.4	77.0	28564	9	ADD46613 Human gen
38	15.4	77.0	335913	5	AAI61371 Soybean 2
39	15.4	77.0	335913	5	AAI61372 Soybean 2
40	15.2	76.0	308	2	AAV75448 Staphyloc
41	15.2	76.0	330	2	AAV75448 Staphyloc
42	15.2	76.0	443	7	AAD50058 Human fat
43	15.2	76.0	468	3	AAFI1349 Aspergill
44	15.2	76.0	472	7	AAD50056 Human fat
45	15.2	76.0	499	6	AAK69101 DNA enco
46	15.2	76.0	499	7	AAD50057 Human fat
47	15.2	76.0	547	6	ABK69131 DNA enco
48	15.2	76.0	548	7	AAD50054 Human fat
49	15.2	76.0	578	3	AAD50053 Human fat
50	15.2	76.0	685	7	AAFI1457 Aspergill
51	15.2	76.0	708	7	AAD50052 Human fat
52	15.2	76.0	785	4	AAK72968 Human imm
53	15.2	76.0	912	7	AAZ51334 Aspergill
54	15.2	76.0	919	6	AAZ51910 Arabidops
55	15.2	76.0	1035	6	ABL91292 Chlamydia
56	15.2	76.0	1107	7	ACA38617 Prokaryot
57	15.2	76.0	1287	5	AAH67093 C. glutami
58	15.2	76.0	1335	4	AAH68053 Corynebac
59	15.2	76.0	1410	6	AAH68052 Corynebac
60	15.2	76.0	2000	6	ABZ17137 Arabidops
61	15.2	76.0	2000	7	ADA68826 Arabidops
62	15.2	76.0	2368	3	AAZ49575 S. aureus
63	15.2	76.0	2385	2	AAQ27956 Cholester
64	15.2	76.0	2608	3	AAZ49578 S. aureus
65	15.2	76.0	2916	8	ADA02611 Mouse Nfk
66	15.2	76.0	2916	9	ADB72349 Mouse Nfk
67	15.2	76.0	2916	9	ADB82943 Mouse Nfk
68	15.2	76.0	3032	5	ABV25508 Human pro
69	15.2	76.0	3324	7	AAD50048 Human fat
70	15.2	76.0	3334	6	ABK35711 CDNA sequ
71	15.2	76.0	3657	7	AAD50047 Human fat
72	15.2	76.0	3663	7	AAD50046 Human fat
73	15.2	76.0	3670	9	ADB75301 Prostate
74	15.2	76.0	3762	7	AAD50064 Human fat
75	15.2	76.0	3768	7	AAD50063 Human fat
76	15.2	76.0	3892	6	AB199776 Mouse Isg
77	15.2	76.0	3892	8	ADA02610 Mouse Nfk
78	15.2	76.0	3892	9	ADB72348 Mouse Nfk
79	15.2	76.0	3892	9	ADB82942 Mouse Nfk
80	15.2	76.0	4884	8	ADA31191 DNA enco
81	15.2	76.0	5327	4	AAAF30428 Vesicular
82	15.2	76.0	5674	4	AB126242 Drosophil
83	15.2	76.0	5881	9	ADD27432 Human adi
84	15.2	76.0	5881	9	ADD27605 Human adi
85	15.2	76.0	6222	8	ADA02758 Human MOR
86	15.2	76.0	6222	9	ADB72496 Human MOR
87	15.2	76.0	6222	9	ADB85238 Human MOR
88	15.2	76.0	6330	8	AAAF30429 Vesicular
89	15.2	76.0	6537	8	ADA02757 Human MOR
90	15.2	76.0	6537	9	ADB72495 Human MOR
91	15.2	76.0	6537	9	ADB85237 Human MOR
92	15.2	76.0	7798	9	ADB31351 Human dia
93	15.2	76.0	8354	4	ABL02619 Drosophil
94	15.2	76.0	14311	2	AAAT38744 Plasmid p
95	15.2	76.0	14311	5	AAAF28116 Vesicular
96	15.2	76.0	14311	5	AAAF28115 Vesicular

97	15.2	76.0	14361	4	ABL02618	AbI02618 Drosophi1	C 170	14.8	74.0	1165	6	ABS54004	AbS54004 cDNA enco
98	15.2	76.0	17173	7	ABZ80234	AbZ80234 Mouse tra	C 171	14.8	74.0	1201	7	ABT33402	AbT33402 NOVA DNA
99	15.2	76.0	36412	9	ABD74383	AbD74383 Mycobacte	C 172	14.8	74.0	1242	7	AAF90254	AAf90254 DNA encod
C 100	15.2	76.0	82993	6	ABX09140	AbX09140 Mycobacte	C 173	14.8	74.0	1242	5	AAST7861	AAst7861 DNA encod
C 101	15.2	76.0	96598	8	ADN02609	AdN02609 Mouse NfK	C 174	14.8	74.0	1242	6	AAD21891	AA221891 Human tum
C 102	15.2	76.0	96598	8	ADN02609	AdN02609 Mouse NfK	C 175	14.8	74.0	1357	9	ADC86072	AdC86072 Human GPC
C 103	15.2	76.0	96598	8	ADN02609	AdN02609 Mouse NfK	C 176	14.8	74.0	1357	9	ADC86070	AdC86070 Human GPC
C 104	15.2	76.0	96598	8	ADN02609	AdN02609 Mouse NfK	C 177	14.8	74.0	1357	9	AAAG62348	AAg62348 cDNA sequ
C 105	15.2	76.0	110000	2	AAK91990_01	ContInnation (2 of	C 178	14.8	74.0	1568	3	AAAG98032	AAg98032 Human col
C 106	15.2	76.0	110000	4	AAI99682_21	Continuation (22 o	C 179	14.8	74.0	1719	7	ACA52604	ACa52604 Prokaryot
C 107	15.2	76.0	326014	6	ABK892926	AbK892926 Human gen	C 180	14.8	74.0	1818	7	ACA31066	ACa31066 Prokaryot
C 108	15.2	76.0	349980	5	AAH68530	AAh68530 c glutam1	C 181	14.8	74.0	1836	4	AAI19464	AAi19464 Probe #93
C 109	15	75.0	593	4	AAI17156	AAi17156 Probe #70	C 182	14.8	74.0	1836	4	ABA64480	ABa64480 Human foe
C 110	15	75.0	593	4	ABA61657	ABa61657 Human foe	C 183	14.8	74.0	1836	4	AAI44657	AAi44657 Probe #13
C 111	15	75.0	593	4	AAI41569	AAi41569 Probe #10	C 184	14.8	74.0	1836	4	ABA31619	ABa31619 Probe #10
C 112	15	75.0	593	4	ABA29315	ABa29315 Probe #77	C 185	14.8	74.0	1836	4	AAK12937	AAk12937 Human bra
C 113	15	75.0	593	4	AAK35852	AAk35852 Human don	C 186	14.8	74.0	1836	4	ABS38231	ABs38231 Human liv
C 114	15	75.0	593	4	AAK09958	AAk09958 Human bra	C 187	14.8	74.0	1836	6	ABS12734	ABs12734 Human gen
C 115	15	75.0	593	4	ABS35562	ABs35562 Human liv	C 188	14.8	74.0	1859	7	ABZ57038	ABz57038 Human m11
C 116	15	75.0	593	6	ABS10085	ABs10085 Human gen	C 189	14.8	74.0	1945	4	ABL16181	ABl16181 Drosophi1
C 117	15	75.0	761	9	ADD17343	AdD17343 DNA (Seg1	C 190	14.8	74.0	1987	4	AAI13905	AAi13905 Probe #38
C 118	15	75.0	1851	9	ADD43533	AdD43533 Mouse ope	C 191	14.8	74.0	1987	4	ABA55617	ABa55617 Human foe
C 119	15	75.0	2000	9	ADD43535	AdD43535 Mouse cDN	C 192	14.8	74.0	1987	4	AAI35272	AAi35272 Probe #39
C 120	14.8	74.0	30	6	ABX69902	AbX69902 Novel He1	C 193	14.8	74.0	1987	4	ABA45139	ABa45139 Human bre
C 121	14.8	74.0	100	7	ACD78987	ACd78987 E. coli K	C 194	14.8	74.0	1987	4	ABA25315	ABa25315 Probe #37
C 122	14.8	74.0	138	7	ACA41455	ACa41455 Prokaryot	C 195	14.8	74.0	1987	4	AAK03845	AAk03845 Human bon
C 123	14.8	74.0	276	4	AAK63634	AAk63634 Human imm	C 196	14.8	74.0	1987	4	ABS28932	ABs28932 Human liv
C 124	14.8	74.0	315	8	ADA299841	AdA299841 DNA encod	C 197	14.8	74.0	1987	5	AAI03774	AAi03774 Probe #37
C 125	14.8	74.0	348	8	ADA299940	AdA299940 DNA encod	C 198	14.8	74.0	1987	5	ABS03870	ABs03870 Human gen
C 126	14.8	74.0	380	5	AAI05027	AAi05027 Probe #50	C 199	14.8	74.0	1987	6	ADA72679	ADa72679 Rice gene
C 127	14.8	74.0	380	5	AAI09998	AAi09998 Probe #99	C 200	14.8	74.0	2000	7	ABS51296	ABs51296 Human REM
C 128	14.8	74.0	380	5	ABS05391	ABs05391 Human gen	C 201	14.8	74.0	2033	6	ABZ14366	ABz14366 Arabidops
C 129	14.8	74.0	380	6	ABS17888	ABs17888 Human gen	C 202	14.8	74.0	2379	2	AAK90784	AAk90784 Mutant DN
C 130	14.8	74.0	414	7	ACA19003	ACa19003 Prokaryot	C 203	14.8	74.0	2737	2	AAK90780	AAk90780 PCR ampli
C 131	14.8	74.0	459	4	AAI10174	AAi10174 Probe #10	C 204	14.8	74.0	2737	2	AAK90785	AAk90785 Deterge
C 132	14.8	74.0	459	4	ABA51807	ABa51807 Human foe	C 205	14.8	74.0	2737	2	AAH57377	AAh57377 Human hea
C 133	14.8	74.0	459	4	AAI31424	AAi31424 Probe #11	C 206	14.8	74.0	2814	4	ABL18346	ABl18346 Drosophi1
C 134	14.8	74.0	459	4	ABA21636	ABa21636 Probe #10	C 207	14.8	74.0	3115	6	ABN87601	ABn87601 Human cop
C 135	14.8	74.0	459	4	AAK00114	AAk00114 Human bra	C 208	14.8	74.0	3153	6	ABN87601	ABn87601 Human cop
C 136	14.8	74.0	459	4	ABS25121	ABs25121 Human liv	C 209	14.8	74.0	3295	9	ADB68998	ADb68998 C. neofor
C 137	14.8	74.0	459	6	ABS00120	ABs00120 Human gen	C 210	14.8	74.0	3869	4	AAK79858	AAk79858 Human imm
C 138	14.8	74.0	540	5	ABS72578	ABs72578 DNA encod	C 211	14.8	74.0	4042	9	ADB07449	ADb07449 Novel cod
C 139	14.8	74.0	546	8	ADA29705	ADa29705 DNA encod	C 212	14.8	74.0	4125	7	ADA70202	ADa70202 Rice gene
C 140	14.8	74.0	577	4	AAI23122	AAi23122 Probe #13	C 213	14.8	74.0	4212	7	ABL16180	ABl16180 Drosophi1
C 141	14.8	74.0	577	4	ABA68215	ABa68215 Human foe	C 214	14.8	74.0	4595	6	ABK63624	ABk63624 Rat seque
C 142	14.8	74.0	577	4	AAI48430	AAi48430 Probe #17	C 215	14.8	74.0	4595	9	ADB58031	ADb58031 Toxicity-
C 143	14.8	74.0	577	4	ABA50225	ABa50225 Human bre	C 216	14.8	74.0	4595	9	ADB52503	ADb52503 Primary r
C 144	14.8	74.0	577	4	ABA35227	ABa35227 Probe #13	C 217	14.8	74.0	5438	5	ABV24305	ABv24305 Human pro
C 145	14.8	74.0	577	4	AAK42351	AAk42351 Human bon	C 218	14.8	74.0	7047	6	AAI49486	AAi49486 Drosophi1
C 146	14.8	74.0	577	4	AAK16596	AAk16596 Human bra	C 219	14.8	74.0	7047	6	ABN85738	ABn85738 Drosophi1
C 147	14.8	74.0	577	4	ABS41960	ABs41960 Human liv	C 220	14.8	74.0	7156	2	AAV15586	AAv15586 Lactobaci
C 148	14.8	74.0	577	5	AAI08777	AAi08777 Probe #87	C 221	14.8	74.0	7468	4	ABL05347	ABl05347 Drosophi1
C 149	14.8	74.0	577	6	ABS16406	ABs16406 Human gen	C 222	14.8	74.0	7566	4	AAI07322	AAi07322 Human rep
C 150	14.8	74.0	635	3	AAAF08438	AAaF08438 Fusarilum	C 223	14.8	74.0	7601	4	AAK81252	AAk81252 Human imm
C 151	14.8	74.0	642	7	ACA31472	ACa31472 Prokaryot	C 224	14.8	74.0	7741	7	ABX97045	ABx97045 Human NOV
C 152	14.8	74.0	784	4	AAK85396	AAk85396 Human imm	C 225	14.8	74.0	8008	8	AAAD51565	AAaD51565 Human str
C 153	14.8	74.0	858	6	ABZ43136	ABz43136 Human GPC	C 226	14.8	74.0	11656	4	ABL05346	ABl05346 Drosophi1
C 154	14.8	74.0	858	6	ABZ43134	ABz43134 Human GPC	C 227	14.8	74.0	13954	2	AAK20641	AAk20641 Polynucle
C 155	14.8	74.0	944	6	ABU92761	ABi92761 G protei	C 228	14.8	74.0	14668	4	ABL28811	ABl28811 Drosophi1
C 156	14.8	74.0	954	4	AAH31607	AAh31607 Human olf	C 229	14.8	74.0	15246	7	ABT13955	ABt13955 Simian vi
C 157	14.8	74.0	954	4	AAH32129	AAh32129 Human olf	C 230	14.8	74.0	18060	4	ABL14126	ABl14126 Drosophi1
C 158	14.8	74.0	957	5	AAAG42395	AAaG42395 Human cDN	C 231	14.8	74.0	19174	4	AAK81253	AAk81253 Human imm
C 159	14.8	74.0	957	5	AAAG42396	AAaG42396 Human cDN	C 232	14.8	74.0	29283	9	ADC87182	ADc87182 Human GPC
C 160	14.8	74.0	957	6	ABK68559	ABk68559 Human DNA	C 233	14.8	74.0	29781	4	ABL28810	ABl28810 Drosophi1
C 161	14.8	74.0	957	6	ABK68557	ABk68557 Human DNA	C 234	14.8	74.0	31203	4	AAK81254	AAk81254 Human imm
C 162	14.8	74.0	957	6	ABK37662	ABk37662 DNA encod	C 235	14.8	74.0	32167	5	ABA20621	ABa20621 Human ner
C 163	14.8	74.0	957	6	ABK37661	ABk37661 DNA encod	C 236	14.8	74.0	33675	6	AAAD31202	AAaD31202 Oryza sat
C 164	14.8	74.0	972	6	ABU92762	ABi92762 G protei	C 237	14.8	74.0	35498	4	ABL20316	ABl20316 Drosophi1
C 165	14.8	74.0	1031	6	ABL92760	ABl92760 G protei	C 238	14.8	74.0	52216	6	AAH28355	AAh28355 Nucleoicd
C 166	14.8	74.0	1051	6	ABL92763	ABl92763 G protei	C 239	14.8	74.0	52216	6	ABL50307	ABl50307 Human mus
C 167	14.8	74.0	1133	7	AAI60086	AAi60086 M. annula	C 240	14.8	74.0	110000	6	ABS5320_1	ABs5320_1 ContInnation (2 of
C 168	14.8	74.0	1165	2	AAV08188	AAv08188 Rabbit in	C 241	14.8	74.0	110000	6	ABX08336_03	ABx08336_03 ContInnation (4 of
C 169	14.8	74.0	1165	2	AAV63232	AAv63232 Rabbit in	C 242	14.8	74.0	118584	9	ADC87623	ADc87623 Human GPC

C 243	14.8	74.0	126512	6	ABN83429	Abn83429 Human ttxa	316	14.2	71.0	165	4	AAK20799	Aak20799 Human bra
244	14.8	74.0	147724	6	ABK83566	Abk83566 Human cdn	317	14.2	71.0	165	4	ABE46714	AbE46714 Human liv
245	14.8	74.0	173810	6	ABN835752	Abn835752 Mouse chr	318	14.2	71.0	165	6	ABE21216	ABE21216 Human gen
246	14.8	74.0	178896	6	ABO88146	Abg88146 Human osc	319	14.2	71.0	170	6	ABX92109	ABX92109 Lung spec
C 247	14.8	74.0	199377	9	ADC35071	Adc35071 Mouse gen	320	14.2	71.0	209	7	ABX43080	Abx43080 Bovine ES
C 248	14.8	74.0	335199	9	ADC24703	Adc24703 Human wll	321	14.2	71.0	218	4	AAH34647	AaH34647 Human col
249	14.8	74.0	341511	6	ABSS5200	Abss5200 Genomic D	322	14.2	71.0	230	2	AAAT2311	AAAT2311 Human gen
250	14.8	74.0	349999	9	ADC87010	Adc87010 Human GPC	323	14.2	71.0	231	3	AAAC28393	AAAC28393 Human sec
C 251	14.4	72.0	356	6	ABL93045	AbL93045 Rat metas	324	14.2	71.0	234	3	AA667889	AA667889 plms rtd
C 252	14.4	72.0	392	8	ACH17198	Ach17198 Human adu	325	14.2	71.0	247	4	AA122630	AA122630 Probe #12
C 253	14.4	72.0	402	4	AA189035	AA189035 Human pol	326	14.2	71.0	247	4	ABA67708	ABa67708 Human toe
254	14.4	72.0	540	6	ABN86824	Abn86824 Streptoco	327	14.2	71.0	247	4	AA147924	AA147924 Probe #16
255	14.4	72.0	576	4	AA523940	AA523940 Human ova	328	14.2	71.0	247	4	ABA49791	ABa49791 Human bre
256	14.4	72.0	576	5	AAH82498	AAh82498 Human ova	329	14.2	71.0	247	4	ABA34769	ABa34769 Probe #13
257	14.4	72.0	582	4	AA524374	AA524374 Human ova	330	14.2	71.0	247	4	AAK41863	AAK41863 Human bon
258	14.4	72.0	592	5	AAH82967	AAh82967 Human ova	331	14.2	71.0	247	4	AAK16122	AAK16122 Human bon
C 259	14.4	72.0	631	9	ADD34944	Add34944 Mouse mlt	332	14.2	71.0	247	4	ABE41464	ABe41464 Human liv
C 260	14.4	72.0	627	9	ADB85498	AdB85498 Toxicity-	333	14.2	71.0	247	5	AA108301	AA108301 Probe #82
C 261	14.4	72.0	1155	6	AA598873	AA598873 Breat cu	334	14.2	71.0	247	6	AB815890	AB815890 Human gen
C 262	14.4	72.0	1155	6	ABO86166	Abg86166 Novel hum	335	14.2	71.0	255	5	AAH82259	AAH82259 Rat diffe
C 263	14.4	72.0	1155	7	ADA11408	Ada11408 Human bre	336	14.2	71.0	283	5	ABV47390	ABv47390 Human pro
C 264	14.4	72.0	1155	9	ADC15381	Adc15381 Human bre	337	14.2	71.0	298	7	ABX42500	ABX42500 Bovine ES
C 265	14.4	72.0	1365	7	ADA70035	Ada70035 Rice gene	C 338	14.2	71.0	300	2	AAV75297	AAV75297 Staphyloc
C 266	14.4	72.0	1452	3	AAAC51127	AAc51127 Arabidops	339	14.2	71.0	329	2	AACT71145	AAcT71145 Single nu
C 267	14.4	72.0	1554	7	ACF72121	Acf72121 Phototrab	C 340	14.2	71.0	329	3	AACT71148	AAcT71148 Single nu
C 268	14.4	72.0	1661	3	AAAC36273	AAc36273 Arabidops	C 341	14.2	71.0	329	3	AACT71148	AAcT71148 Single nu
C 269	14.4	72.0	1722	7	ACA31375	ACA31375 Prokaryot	C 342	14.2	71.0	329	3	AACT71148	AAcT71148 Single nu
C 270	14.4	72.0	1953	4	AAE61012	AAe61012 P. putida	C 343	14.2	71.0	330	4	AA527147	AA527147 CDNA enco
C 271	14.4	72.0	2000	7	ACA36218	ACA36218 Prokaryot	C 344	14.2	71.0	330	4	AA530225	AA530225 DNA enco
C 272	14.4	72.0	2295	7	ADA73213	Ada73213 Rice gene	C 345	14.2	71.0	330	4	AA535560	AA535560 Human car
C 273	14.4	72.0	2433	2	AAO89728	AAo89728 Prokaryot	C 346	14.2	71.0	330	7	ACD01480	ACD01480 Human pol
C 274	14.4	72.0	2437	2	ABT32215	Abt32215 Plasmodu	C 347	14.2	71.0	330	9	ADB93325	ADB93325 Human CDN
C 275	14.4	72.0	2492	4	ABL23842	AbL23842 Drosophi1	C 348	14.2	71.0	330	9	ADB45639	ADB45639 Human car
C 276	14.4	72.0	2516	4	ABLT5452	ABlT5452 Drosophi1	C 349	14.2	71.0	343	5	ABE11960	ABe11960 Human ner
C 277	14.4	72.0	2554	3	AAZ29341	AAz29341 Tomato be	C 350	14.2	71.0	343	5	ABE11961	ABe11961 Human ner
C 278	14.4	72.0	2663	2	AAO89728	AAo89728 Plasmodu	C 351	14.2	71.0	348	7	ABE240462	ABe240462 N. gonorr
C 279	14.4	72.0	2871	2	ABL20186	ABl20186 Drosophi1	C 352	14.2	71.0	361	5	ABA11559	ABa11559 Human ner
C 280	14.4	72.0	2871	4	ABL22812	ABl22812 Drosophi1	C 353	14.2	71.0	379	6	ABN92985	ABn92985 Staphyloc
C 281	14.4	72.0	2937	4	ABL21888	ABl21888 Drosophi1	C 354	14.2	71.0	379	6	ABN22923	ABn22923 Human ONF
C 282	14.4	72.0	3005	7	AAK95954	AAK95954 MDDT re1a	C 355	14.2	71.0	381	7	ACA48990	ACA48990 Prokaryot
C 283	14.4	72.0	4062	2	AAAT13347	AAaT13347 C3G prote	C 356	14.2	71.0	382	7	ABX46334	ABx46334 Bovine ES
C 284	14.4	72.0	4094	6	AA594896	AA594896 Human DNA	C 357	14.2	71.0	384	7	ACA50945	ACA50945 Prokaryot
C 285	14.4	72.0	9235	2	AAAX13131	AAx13131 Enterococ	C 358	14.2	71.0	384	7	ACA51451	ACA51451 Prokaryot
C 286	14.4	72.0	9235	6	AB598926	AB598926 Enterococ	C 359	14.2	71.0	395	2	AAH83730	AAH83730 DNA enco
C 287	14.4	72.0	11337	4	AAU36286	AAU36286 Human mus	C 360	14.2	71.0	401	4	AA115814	AA115814 Probe #57
C 288	14.4	72.0	11337	7	ABX59274	ABx59274 CDNA enco	C 361	14.2	71.0	401	4	ABAS8027	ABAS8027 Human toe
C 289	14.4	72.0	12776	4	ABL11894	ABl11894 Drosophi1	C 362	14.2	71.0	401	4	AA137628	AA137628 Probe #63
C 290	14.4	72.0	12878	4	ABL13212	ABl13212 Drosophi1	C 363	14.2	71.0	401	4	ABA27288	ABa27288 Probe #57
C 291	14.4	72.0	16656	2	AAAT30787	AAaT30787 Alphasvitu	C 364	14.2	71.0	401	4	AAK31752	AAK31752 Human bon
C 292	14.4	72.0	24079	4	AAK71153	AAK71153 Human imm	C 365	14.2	71.0	401	4	AAK06101	AAK06101 Human bra
C 293	14.4	72.0	24079	7	ABZ67330	ABz67330 Human sec	C 366	14.2	71.0	401	4	AB531436	AB531436 Human liv
C 294	14.4	72.0	24079	7	ABZ73732	ABz73732 Seceted	C 367	14.2	71.0	401	6	AB506508	AB506508 Human gen
C 295	14.4	72.0	24079	7	ADA98437	ADa98437 Human sec	C 368	14.2	71.0	433	7	ACH439256	ACH439256 Bovine ES
C 296	14.4	72.0	24079	9	ADC20606	Adc20606 Human sec	C 369	14.2	71.0	437	8	ACH44707	ACH44707 Human toe
297	14.4	72.0	50000	9	ADC60733	Adc60733 Human sli	C 370	14.2	71.0	465	3	AAAC37114	AAAC37114 Arabidops
298	14.4	72.0	50341	3	AAV26274	AAv26274 DNA seque	C 371	14.2	71.0	466	4	AA110463	AA110463 Probe #39
299	14.4	72.0	50341	3	AAZ29519	AAz29519 US snutcl	C 372	14.2	71.0	466	4	ABAS2109	ABAS2109 Human toe
300	14.4	72.0	52297	2	AAAT51411	AAaT51411 Mycobacte	C 373	14.2	71.0	466	4	AA131716	AA131716 Probe #40
301	14.4	72.0	52298	2	AAQ47357	AAq47357 US mycoba	C 374	14.2	71.0	466	4	ABA21919	ABa21919 Probe #38
C 302	14.4	72.0	67571	7	AAAD53224_5	Continuaction (6 of	C 375	14.2	71.0	466	4	AAK25834	AAK25834 Human bon
C 303	14.4	72.0	75798	7	ABX77212	Abx77212 Mouse urt	C 376	14.2	71.0	466	4	AAK00391	AAK00391 Human bra
304	14.4	72.0	110000	7	ACF67367_53	Continuaction (54 o	C 377	14.2	71.0	466	4	AB525422	AB525422 Human liv
305	14.4	72.0	110000	7	ACF65367_5	Continuaction (6 of	C 378	14.2	71.0	466	5	AA100399	AA100399 Probe #39
C 306	14.2	71.0	23	9	ADB88615	AdB88615 Frizlied-	C 379	14.2	71.0	466	6	AB500415	AB500415 Human gen
307	14.2	71.0	33	6	AAZ47523	AAz47523 Human rib	C 380	14.2	71.0	473	6	ABAS6963	ABAS6963 Human toe
308	14.2	71.0	36	3	AAZ290660	AAz290660 Nostoc Ns	C 381	14.2	71.0	473	6	AB505259	AB505259 Human gen
C 309	14.2	71.0	65	6	ABN51703	ABn51703 Mouse sp1	C 382	14.2	71.0	477	8	ACH24857	ACH24857 Human adu
C 310	14.2	71.0	160	6	ABAS69525	ABa69525 Human foe	C 383	14.2	71.0	482	8	ACH46337	ACH46337 Human inf
C 311	14.2	71.0	160	6	AB517760	ABs17760 Human foe	C 384	14.2	71.0	484	4	AA113401	AA113401 Probe #33
312	14.2	71.0	165	4	ABAA72373	ABaA72373 Human foe	C 385	14.2	71.0	484	4	ABAS5099	ABAS5099 Human toe
313	14.2	71.0	165	4	AA152785	AAa152785 Probe #21	C 386	14.2	71.0	484	4	AA134756	AA134756 Probe #34
314	14.2	71.0	165	4	ABA38182	ABa38182 Probe #16	C 387	14.2	71.0	484	4	ABA44645	ABa44645 Human bre
315	14.2	71.0	165	4	AAK46949	AAK46949 Human bon	C 388	14.2	71.0	484	4	ABA24648	ABa24648 Probe #33

389	14.2	71.0	484	4	AAK28813	Aak28813	Human bon	C 462	14.2	71.0	1452	6	AB214597	Ab214597	Arabidops
390	14.2	71.0	484	4	AAK03362	Aak03362	Human bra	C 463	14.2	71.0	1470	8	AB213991	Ab213991	Arabidops
391	14.2	71.0	484	4	AB284422	Ab284422	Human liv	C 464	14.2	71.0	1477	6	ADA09909	Ada09909	Human rec
392	14.2	71.0	484	5	AA103289	Aa103289	Probe #32	C 465	14.2	71.0	1481	5	AA580066	Aa580066	DNA encod
393	14.2	71.0	486	6	AB503343	Ab503343	Human gen	C 466	14.2	71.0	1536	6	ABN70708	Abn70708	Streptoco
394	14.2	71.0	486	4	AB547502	Ab547502	Human foe	C 467	14.2	71.0	1574	6	AB515109	Ab515109	Human EDA
395	14.2	71.0	486	4	AA137055	Aa137055	Probe #57	C 468	14.2	71.0	1574	8	ACD07895	Act07895	C. neofor
396	14.2	71.0	486	4	ABA26990	Ab26990	Probe #54	C 469	14.2	71.0	1591	9	ADB69393	Ad69393	C. neofor
397	14.2	71.0	486	4	AAK31148	Aak31148	Human bon	C 470	14.2	71.0	1592	3	AACT3871	Aac73871	Human sec
398	14.2	71.0	486	4	AAK05545	Aak05545	Human bra	C 471	14.2	71.0	1604	4	AAK81317	Aak81317	Human imm
399	14.2	71.0	486	4	AB530828	Ab530828	Human liv	C 472	14.2	71.0	1626	6	ABN70294	Abn70294	Streptoco
400	14.2	71.0	486	6	AB505900	Ab505900	Human gen	C 473	14.2	71.0	1626	6	ABN66277	Abn66277	Streptoco
401	14.2	71.0	493	4	AAH11545	Aah11545	Human CDN	C 474	14.2	71.0	1630	3	AAAC5415	Aac5415	Arabidops
402	14.2	71.0	525	4	AAHS3240	Aahs3240	S. epider	C 475	14.2	71.0	1635	3	AAAC5245	Aac5245	Arabidops
403	14.2	71.0	548	4	ABA59832	Ab59832	Human foe	C 476	14.2	71.0	1641	3	AAAC59785	Aac59785	Human sec
404	14.2	71.0	548	4	AA139704	Aa139704	Probe #83	C 477	14.2	71.0	1642	4	ABLI0295	Abli0295	Drosophil
405	14.2	71.0	548	4	ABA28303	Ab28303	Probe #67	C 478	14.2	71.0	1651	8	ACD98388	Act98388	Human gph
406	14.2	71.0	548	4	AAK33981	Aak33981	Human bon	C 479	14.2	71.0	1651	8	ADD30267	Ad30267	Plant yle
407	14.2	71.0	548	4	AAK08105	Aak08105	Human bra	C 480	14.2	71.0	1653	7	ACA339754	Act339754	Prokaryot
408	14.2	71.0	548	4	AB533784	Ab533784	Human liv	C 481	14.2	71.0	1661	6	ABN68861	Abn68861	Streptoco
409	14.2	71.0	548	6	AB508782	Ab508782	Human gen	C 482	14.2	71.0	1707	6	ABN91468	Abn91468	Staphyloc
410	14.2	71.0	555	9	ADC91108	Adc91108	E. fecali	C 483	14.2	71.0	1725	7	ACA43064	Act43064	Prokaryot
411	14.2	71.0	562	9	ADD33014	Ad333014	Human mit	C 484	14.2	71.0	1790	7	AAK63505	Aak63505	Human CDN
412	14.2	71.0	571	5	ABV35222	Abv35222	Human pro	C 485	14.2	71.0	1790	10	ADE76985	Ad76985	Nucleotid
413	14.2	71.0	586	6	ABN66037	Abn66037	Human can	C 486	14.2	71.0	1816	4	AAE25489	Aae25489	Rice gene
414	14.2	71.0	592	4	AAK79691	Aak79691	Human pro	C 487	14.2	71.0	1820	7	ADA70905	Ada70905	Rice gene
415	14.2	71.0	600	5	ABV14127	Abv14127	Human imm	C 488	14.2	71.0	1841	7	AB236174	Ab236174	Human sec
416	14.2	71.0	629	3	AAAF09337	Aaf09337	Fusarium	C 489	14.2	71.0	1843	2	AAO14798	Aa14798	Ro/SSA au
417	14.2	71.0	657	4	ABLI5187	Abli5187	Drosophil	C 490	14.2	71.0	1845	3	AAZ38967	Aaz38967	Escherich
418	14.2	71.0	657	4	ABLI22877	Abli22877	Drosophil	C 491	14.2	71.0	1879	6	ABL59530	Ab59530	Human gph
419	14.2	71.0	693	7	AB272013	Ab272013	Human NAD	C 492	14.2	71.0	1883	9	ADD18526	Ad18526	Human pro
420	14.2	71.0	700	4	ABLI03751	Abli03751	Drosophil	C 493	14.2	71.0	1890	6	ABK84052	Abk84052	Human CDN
421	14.2	71.0	721	2	AAAT35105	Aat35105	Down-regu	C 494	14.2	71.0	1899	7	ACA21814	Act21814	Prokaryot
422	14.2	71.0	743	8	ABLI51017	Abli51017	Human EDA	C 495	14.2	71.0	1902	6	ABO70018	Ab70018	Listeria
423	14.2	71.0	743	8	ACD07993	Act07993	DNA encod	C 496	14.2	71.0	1902	6	ABO68241	Ab68241	Listeria
424	14.2	71.0	744	6	ABK35418	Abk35418	Human CDN	C 497	14.2	71.0	1903	4	AAH16546	Aah16546	Human CDN
425	14.2	71.0	771	4	AAAC53200	Aac53200	Arabidops	C 498	14.2	71.0	1908	7	ACA26970	Act26970	Prokaryot
426	14.2	71.0	776	6	ABN66321	Abn66321	Streptoco	C 499	14.2	71.0	1953	3	AAA09499	Aa09499	Plant-opt
427	14.2	71.0	776	4	AAI95828	Aai95828	Human neu	C 500	14.2	71.0	1962	3	AAZ51637	Aaz51637	Delta-ent
428	14.2	71.0	789	3	AAAC53161	Aac53161	Arabidops	C 501	14.2	71.0	1962	7	ABX95181	Abx95181	DNA encod
429	14.2	71.0	801	2	AAH07525	Aah07525	Human CDN	C 502	14.2	71.0	1962	9	AAD61785	Aad61785	Bacillus
430	14.2	71.0	807	4	AAH39942	Aah39942	Gastric c	C 503	14.2	71.0	1989	9	AAZ51638	Aaz51638	Bacillus
431	14.2	71.0	830	2	AAI216633	Aai216633	Human gen	C 504	14.2	71.0	1989	6	AB565041	Ab565041	Invertebr
432	14.2	71.0	870	7	ACA39785	Act39785	Prokaryot	C 505	14.2	71.0	1989	7	ABX95182	Abx95182	DNA encod
433	14.2	71.0	896	5	ABAI5574	Abai5574	Human ner	C 506	14.2	71.0	1989	9	AAD61786	Aad61786	Bacillus
434	14.2	71.0	915	7	ACA51998	Act51998	Prokaryot	C 507	14.2	71.0	2000	2	AAH81529	Aah81529	Plantized
435	14.2	71.0	917	3	AAAC36836	Aac36836	Arabidops	C 508	14.2	71.0	2000	7	ADA73099	Ad73099	Rice gene
436	14.2	71.0	919	6	ABUS8803	Abi58803	Fungal me	C 509	14.2	71.0	2000	7	ADA72245	Ad72245	Rice gene
437	14.2	71.0	930	7	ACA39461	Act39461	Prokaryot	C 510	14.2	71.0	2000	7	ADA72245	Ad72245	Rice gene
438	14.2	71.0	1025	6	AAAD1879	Ad1879	Lactobact	C 511	14.2	71.0	2002	2	AAV43605	Aav43605	Human sec
439	14.2	71.0	1059	6	ABLI40697	Abli40697	Human big	C 512	14.2	71.0	2020	2	AAQ53209	Aaq53209	Cyclin D3
440	14.2	71.0	1095	4	ABLI22477	Abli22477	Drosophil	C 513	14.2	71.0	2022	2	AAQ31877	Aaq31877	DNA seque
441	14.2	71.0	1146	7	ACA32818	Act32818	Prokaryot	C 514	14.2	71.0	2050	2	AAH81530	Aah81530	Human pan
442	14.2	71.0	1148	9	ADD30958	Ad30958	Plant yle	C 515	14.2	71.0	2109	3	AAAC98855	Aac98855	Human pan
443	14.2	71.0	1176	6	ABLI51020	Abli51020	Human EDA	C 516	14.2	71.0	2110	9	ADB62589	Ad62589	Human CDN
444	14.2	71.0	1176	8	ACD07906	Act07906	DNA encod	C 517	14.2	71.0	2152	2	AAQ70732	Aaq70732	TATA-bind
445	14.2	71.0	1192	4	AAAC36066	Aac36066	Arabidops	C 518	14.2	71.0	2152	2	AAAT42218	Aat42218	Human TAT
446	14.2	71.0	1213	4	AAAF63802	Aaf63802	Human sec	C 519	14.2	71.0	2152	2	AAAT79594	Aat79594	TATA-bind
447	14.2	71.0	1214	5	AA529071	Aa529071	CDNA enco	C 520	14.2	71.0	2200	3	AAAC77823	Aac77823	Human can
448	14.2	71.0	1214	6	AB568211	Ab568211	CDNA enco	C 521	14.2	71.0	2211	5	AAAC91228	Aac91228	Human S-8
449	14.2	71.0	1214	6	ADC52205	Adc52205	Human CDN	C 522	14.2	71.0	2215	7	ACA45338	Act45338	Prokaryot
450	14.2	71.0	1225	2	AAAT33325	Aaat33325	American	C 523	14.2	71.0	2232	5	AA594081	Aa594081	DNA encod
451	14.2	71.0	1248	7	ACA48457	Act48457	Prokaryot	C 524	14.2	71.0	2259	2	AAAT30686	Aat30686	Kaposi's
452	14.2	71.0	1266	7	ABX06878	Abx06878	S. pneumo	C 525	14.2	71.0	2259	2	AAAT16816	Aat16816	Kaposi's
453	14.2	71.0	1281	7	ACA39944	Act39944	Prokaryot	C 526	14.2	71.0	2274	7	ACA43678	Act43678	Prokaryot
454	14.2	71.0	1287	7	ACA31014	Act31014	Prokaryot	C 527	14.2	71.0	2305	7	ADB89728	Ad89728	Staphyloc
455	14.2	71.0	1290	6	AA529143	Aa529143	CDNA enco	C 528	14.2	71.0	2321	9	ADB56281	Ad56281	Human gen
456	14.2	71.0	1290	6	AB568283	Ab568283	CDNA enco	C 529	14.2	71.0	2321	9	ADB56281	Ad56281	Human gen
457	14.2	71.0	1290	3	ADC52277	Act52277	Human CDN	C 530	14.2	71.0	2344	2	AAQ33394	Aaq33394	R496L ASM
458	14.2	71.0	1305	3	AAAC52177	Aac52177	Human CDN	C 531	14.2	71.0	2344	8	AAAT95066	Aat95066	CDNA enco
459	14.2	71.0	1312	4	ABLI08777	Abli08777	Drosophil	C 532	14.2	71.0	2344	8	ACD28698	Act28698	CDNA enco
460	14.2	71.0	1365	5	AAAC40275	Aac40275	Arabidops	C 533	14.2	71.0	2346	2	AAAT95068	Aat95068	CDNA enco
461	14.2	71.0	1452	5	AA571632	Aa571632	DNA encod	C 534	14.2	71.0	2347	2	AAQ33390	Aaq33390	ASM CDNA,

535	14.2	71.0	2347	2	AAQ33395	608	14.2	71.0	3222	4	ABL05253	Abi05253	Drosophila
536	14.2	71.0	2347	2	AAQ33393	C 609	14.2	71.0	3243	6	AAD41265	Aad41265	Human ade
537	14.2	71.0	2347	2	AAAT95067	C 610	14.2	71.0	3262	7	AAD54585	Aad54585	Human LIM
538	14.2	71.0	2347	2	AAAT95063	C 611	14.2	71.0	3262	9	ADC77694	Adc77694	Human 161
539	14.2	71.0	2347	2	AAAT95065	C 612	14.2	71.0	3265	3	AAZ65095	Aaz65095	Membrane
540	14.2	71.0	2347	8	ACD28667	C 613	14.2	71.0	3265	4	AAZ64603	Aaz64603	Human DNA
541	14.2	71.0	2347	8	ACD28667	C 614	14.2	71.0	3265	5	AAAT92092	Aat92092	Human PRO
542	14.2	71.0	2347	8	ACD28667	C 615	14.2	71.0	3265	5	AAAT92092	Aat92092	Human PRO
543	14.2	71.0	2362	1	AAN91636	C 616	14.2	71.0	3265	6	ABT74412	Abt74412	Human PRO
544	14.2	71.0	2362	9	ADBS6925	C 617	14.2	71.0	3265	7	ABX78656	Abx78656	Human PRO
545	14.2	71.0	2370	9	ACF73789	C 618	14.2	71.0	3265	7	ACA75628	Ac75628	Human PRO
546	14.2	71.0	2373	6	ABL67219	C 619	14.2	71.0	3265	7	ACA71108	Ac71108	Human sec
547	14.2	71.0	2373	6	ABZ35634	C 620	14.2	71.0	3265	7	ACB87636	Ac87636	Human sec
548	14.2	71.0	2373	6	ABN95867	C 621	14.2	71.0	3265	7	ACB87022	Ac87022	Human sec
549	14.2	71.0	2429	8	ABN63344	C 622	14.2	71.0	3265	7	ACD04195	Ac04195	Human sec
550	14.2	71.0	2429	8	ACH03893	C 623	14.2	71.0	3265	7	ABX77953	Abx77953	Human PRO
551	14.2	71.0	2437	4	AAH16311	C 624	14.2	71.0	3265	7	ABX80365	Abx80365	Human PRO
552	14.2	71.0	2451	2	AAQ14235	C 625	14.2	71.0	3265	7	ACA69271	Ac69271	Human CDN
553	14.2	71.0	2456	7	AAD54586	C 626	14.2	71.0	3265	7	ACA69526	Ac69526	CDNA enco
554	14.2	71.0	2468	5	AAS69788	C 627	14.2	71.0	3265	7	ACA90371	Ac90371	Human PRO
555	14.2	71.0	2485	6	ABZ35397	C 628	14.2	71.0	3265	7	ACC89478	Acc89478	Human sec
556	14.2	71.0	2485	9	ADBS3859	C 629	14.2	71.0	3265	7	ABX90342	Abx90342	Human sec
557	14.2	71.0	2487	9	ACA45126	C 630	14.2	71.0	3265	7	ACA98269	Ac98269	Human sec
558	14.2	71.0	2549	2	AAQ26401	C 631	14.2	71.0	3265	7	ACA93911	Ac93911	Human sec
559	14.2	71.0	2562	5	AAS60002	C 632	14.2	71.0	3265	7	ACD15304	Ac15304	Human sec
560	14.2	71.0	2570	3	AACT6054	C 633	14.2	71.0	3265	7	ACD08891	Ac08891	Human sec
561	14.2	71.0	2558	5	ABV28967	C 634	14.2	71.0	3265	7	ACD96811	Ac96811	Human sec
562	14.2	71.0	2558	5	ABV23126	C 635	14.2	71.0	3265	7	ACF15532	Ac15532	Human sec
563	14.2	71.0	2679	7	ABT15153	C 636	14.2	71.0	3265	7	ABX64188	Abx64188	CDNA enco
564	14.2	71.0	2684	9	ADBS4829	C 637	14.2	71.0	3265	7	ACA72899	Ac72899	Human PRO
565	14.2	71.0	2706	6	ABN93089	C 638	14.2	71.0	3265	7	ACD03071	Ac03071	Human PRO
566	14.2	71.0	2716	4	ABL22886	C 639	14.2	71.0	3265	7	ACD01886	Ac01886	Novel hum
567	14.2	71.0	2718	9	ADC14279	C 640	14.2	71.0	3265	7	ACA92078	Ac92078	Novel hum
568	14.2	71.0	2751	8	ABZ59766	C 641	14.2	71.0	3265	7	ACA89503	Ac89503	CDNA enco
569	14.2	71.0	2751	8	ADA37187	C 642	14.2	71.0	3265	7	ACA73513	Ac73513	Human sec
570	14.2	71.0	2751	8	ADCS1726	C 643	14.2	71.0	3265	7	ACA05828	Ac05828	Human sec
571	14.2	71.0	2767	4	ABL22876	C 644	14.2	71.0	3265	7	ACA66662	Ac66662	CDNA enco
572	14.2	71.0	2771	7	ABZ42840	C 645	14.2	71.0	3265	7	ACA64410	Ac64410	Novel hum
573	14.2	71.0	2782	4	AAZ5495	C 646	14.2	71.0	3265	7	ACA91198	Ac91198	Novel hum
574	14.2	71.0	2816	4	ABL15186	C 647	14.2	71.0	3265	7	ACD81575	Ac81575	Human CDN
575	14.2	71.0	2824	4	ABL03750	C 648	14.2	71.0	3265	7	ACF22037	Ac22037	Human sec
576	14.2	71.0	2834	6	ABL03750	C 649	14.2	71.0	3265	7	ACF19623	Ac19623	Human sec
577	14.2	71.0	2834	6	ABL66293	C 650	14.2	71.0	3265	7	ACD21911	Ac21911	Human sec
578	14.2	71.0	2834	6	ABL64976	C 651	14.2	71.0	3265	7	ACF13076	Ac13076	Human sec
579	14.2	71.0	2834	6	ABL65641	C 652	14.2	71.0	3265	7	ACD25179	Ac25179	Human sec
580	14.2	71.0	2834	6	ABX92203	C 653	14.2	71.0	3265	7	ACF00228	Ac00228	Human sec
581	14.2	71.0	2834	6	ADD71005	C 654	14.2	71.0	3265	7	ACA60397	Ac60397	Novel hum
582	14.2	71.0	2841	7	ABT17669	C 655	14.2	71.0	3265	7	ACA72285	Ac72285	Novel hum
583	14.2	71.0	2841	7	ABT17669	C 656	14.2	71.0	3265	7	ACD04809	Ac04809	Novel hum
584	14.2	71.0	2852	7	ADA53545	C 657	14.2	71.0	3265	7	ACD18270	Ac18270	Human sec
585	14.2	71.0	2852	8	ACH03950	C 658	14.2	71.0	3265	7	ACD08277	Ac08277	Human sec
586	14.2	71.0	2867	8	ABX04186	C 659	14.2	71.0	3265	7	ACA88711	Ac88711	Novel hum
587	14.2	71.0	3043	6	AAD35923	C 660	14.2	71.0	3265	7	ACA70153	Ac70153	Human sec
588	14.2	71.0	3043	6	ABX56624	C 661	14.2	71.0	3265	7	ACD12375	Ac12375	Novel hum
589	14.2	71.0	3044	3	AAZ51656	C 662	14.2	71.0	3265	7	ACC74290	Ac74290	Human sec
590	14.2	71.0	3044	3	ABX95200	C 663	14.2	71.0	3265	7	ACD15918	Ac15918	Human sec
591	14.2	71.0	3044	9	AAH61804	C 664	14.2	71.0	3265	7	ACD25486	Ac25486	Human sec
592	14.2	71.0	3135	4	AAH54194	C 665	14.2	71.0	3265	7	ACD17963	Ac17963	Human sec
593	14.2	71.0	3135	4	ABL22476	C 666	14.2	71.0	3265	7	ACB88250	Ac88250	Human sec
594	14.2	71.0	3169	8	ADA10935	C 667	14.2	71.0	3265	7	ACD21604	Ac21604	Human sec
595	14.2	71.0	3181	6	AAD35925	C 668	14.2	71.0	3265	7	ACD18671	Ac18671	Human sec
596	14.2	71.0	3181	7	ABX56626	C 669	14.2	71.0	3265	7	ACA58844	Ac58844	CDNA enco
597	14.2	71.0	3195	6	ABA91323	C 670	14.2	71.0	3265	7	ABX98281	Abx98281	Human CDN
598	14.2	71.0	3196	9	AAD59152	C 671	14.2	71.0	3265	7	ACD14032	Ac14032	Human PRO
599	14.2	71.0	3198	8	ACD39241	C 672	14.2	71.0	3265	7	ACD09812	Ac09812	Human sec
600	14.2	71.0	3198	8	ACD39240	C 673	14.2	71.0	3265	7	ACB88557	Ac88557	Human sec
601	14.2	71.0	3198	8	ACD39243	C 674	14.2	71.0	3265	7	ACD21297	Ac21297	Human sec
602	14.2	71.0	3198	8	ACD39242	C 675	14.2	71.0	3265	7	ABX75669	Abx75669	Human CDN
603	14.2	71.0	3199	4	ABN09217	C 676	14.2	71.0	3265	7	ACA64020	Ac64020	CDNA enco
604	14.2	71.0	3204	9	ABZ59770	C 677	14.2	71.0	3265	7	ABX97872	Abx97872	Human PRO
605	14.2	71.0	3204	9	AAD58494	C 678	14.2	71.0	3265	7	ACA97348	Ac97348	Human PRO
606	14.2	71.0	3207	6	ABZ35549	C 679	14.2	71.0	3265	7	ACA57811	Ac57811	Human PRO
607	14.2	71.0	3218	6	ABQ81558	C 680	14.2	71.0	3265	7	ACD14339	Ac14339	Human PRO

c 681 14.2 71.0 3265 7 ACC91122 Human sec
c 682 14.2 71.0 3265 7 ACC88864 Human sec
c 683 14.2 71.0 3265 7 ACC07061 Human PRO
c 684 14.2 71.0 3265 7 ACA67512 Human PRO
c 685 14.2 71.0 3265 7 ACC81567 Human sec
c 686 14.2 71.0 3265 7 ACA91284 CDNA enco
c 687 14.2 71.0 3265 7 ACC89110 Human sec
c 688 14.2 71.0 3265 7 ACC86527 Human sec
c 689 14.2 71.0 3265 7 ACC89785 Human sec
c 690 14.2 71.0 3265 7 ACC92964 Human sec
c 691 14.2 71.0 3265 7 ABX80869 Human sec
c 692 14.2 71.0 3265 7 ACA72592 Human PRO
c 693 14.2 71.0 3265 7 ACA89110 Human sec
c 694 14.2 71.0 3265 7 ACA96984 Human sec
c 695 14.2 71.0 3265 7 ACA96989 Novel hum
c 696 14.2 71.0 3265 7 ACA90985 Novel hum
c 697 14.2 71.0 3265 7 ACA70767 Human sec
c 698 14.2 71.0 3265 7 ACA95277 Novel hum
c 699 14.2 71.0 3265 7 ACD44378 CDNA enco
c 700 14.2 71.0 3265 7 ACC86220 Human sec
c 701 14.2 71.0 3265 7 ACD45183 Human sec
c 702 14.2 71.0 3265 7 ACC90092 Human sec
c 703 14.2 71.0 3265 7 ACD12700 Human sec
c 704 14.2 71.0 3265 7 ACF19930 Human sec
c 705 14.2 71.0 3265 7 ABX76874 Human PRO
c 706 14.2 71.0 3265 7 ACA73206 Human PRO
c 707 14.2 71.0 3265 7 ACA68749 Novel hum
c 708 14.2 71.0 3265 7 ACA74593 CDNA enco
c 709 14.2 71.0 3265 7 ACA70460 Human sec
c 710 14.2 71.0 3265 7 ACD14646 Human PRO
c 711 14.2 71.0 3265 7 ACA93731 Human CDN
c 712 14.2 71.0 3265 7 ACA68318 Novel hum
c 713 14.2 71.0 3265 7 ABX98783 Novel hum
c 714 14.2 71.0 3265 7 ACA67305 CDNA enco
c 715 14.2 71.0 3265 7 ACC81260 Human sec
c 716 14.2 71.0 3265 7 ACA95584 Novel hum
c 717 14.2 71.0 3265 7 ACD04502 Novel hum
c 718 14.2 71.0 3265 7 ACC87943 Human sec
c 719 14.2 71.0 3265 7 ACF12605 Human sec
c 720 14.2 71.0 3265 7 ACH66278 Novel hum
c 721 14.2 71.0 3265 7 ABX79549 Human sec
c 722 14.2 71.0 3265 7 ACA96320 Human PRO
c 723 14.2 71.0 3265 7 ACA65094 Human PRO
c 724 14.2 71.0 3265 7 ACA73820 Human sec
c 725 14.2 71.0 3265 7 ACA74232 Novel hum
c 726 14.2 71.0 3265 7 ACA96637 Human PRO
c 727 14.2 71.0 3265 7 ACD10733 CDNA enco
c 728 14.2 71.0 3265 7 ACC91429 Human sec
c 729 14.2 71.0 3265 7 ACA93570 Novel hum
c 730 14.2 71.0 3265 7 ACD02764 CDNA enco
c 731 14.2 71.0 3265 7 ACC87329 Human sec
c 732 14.2 71.0 3265 7 ACC85913 Human sec
c 733 14.2 71.0 3265 7 ABX81252 Novel hum
c 734 14.2 71.0 3265 7 ACA65401 Human PRO
c 735 14.2 71.0 3265 7 ACA94218 Human sec
c 736 14.2 71.0 3265 7 ACA97962 Human PRO
c 737 14.2 71.0 3265 7 ACA91464 Novel hum
c 738 14.2 71.0 3265 7 ACA90678 Novel hum
c 739 14.2 71.0 3265 7 ACD16225 Human sec
c 740 14.2 71.0 3265 7 ACD17386 Human sec
c 741 14.2 71.0 3265 7 ACC92043 Human sec
c 742 14.2 71.0 3265 7 ACD02332 Novel hum
c 743 14.2 71.0 3265 7 ACA74900 CDNA enco
c 744 14.2 71.0 3265 7 ACA91771 Human PRO
c 745 14.2 71.0 3265 7 ACA89323 Novel hum
c 746 14.2 71.0 3265 7 ACA71415 Human sec
c 747 14.2 71.0 3265 7 ACC90815 Human sec
c 748 14.2 71.0 3265 7 ACA65825 CDNA enco
c 749 14.2 71.0 3265 7 ACA68960 Novel hum
c 750 14.2 71.0 3265 7 ACA93068 Novel hum
c 751 14.2 71.0 3265 7 ACA94970 CDNA enco
c 752 14.2 71.0 3265 7 ACD16532 Human sec
c 753 14.2 71.0 3265 7 ACD15611 Human sec

c 754 14.2 71.0 3265 7 ACA99482 Human PRO
c 755 14.2 71.0 3265 7 ABX17152 Human PRO
c 756 14.2 71.0 3265 7 ABX16714 Human CDN
c 757 14.2 71.0 3265 7 ACA68007 Novel hum
c 758 14.2 71.0 3265 8 ACA63407 CDNA enco
c 759 14.2 71.0 3265 8 ACA97655 Human PRO
c 760 14.2 71.0 3265 8 ACA99104 Novel hum
c 761 14.2 71.0 3265 8 ACC91736 Human sec
c 762 14.2 71.0 3265 8 ACD11147 Novel hum
c 763 14.2 71.0 3265 8 ACD14997 Human sec
c 764 14.2 71.0 3265 8 ACA88456 Human sec
c 765 14.2 71.0 3265 8 ACD81963 CDNA enco
c 766 14.2 71.0 3265 8 ACD11761 Human sec
c 767 14.2 71.0 3265 8 ACC95890 Human sec
c 768 14.2 71.0 3265 8 ACF16453 Human sec
c 769 14.2 71.0 3265 8 ACF02571 Human sec
c 770 14.2 71.0 3265 8 ACF02878 Human sec
c 771 14.2 71.0 3265 8 ACF21465 Human sec
c 772 14.2 71.0 3265 8 ACF10149 Human sec
c 773 14.2 71.0 3265 8 ACF78042 Human sec
c 774 14.2 71.0 3265 8 ACD46747 Human sec
c 775 14.2 71.0 3265 8 ACD49510 Human sec
c 776 14.2 71.0 3265 8 ACF28277 Human sec
c 777 14.2 71.0 3265 8 ACD88667 Human sec
c 778 14.2 71.0 3265 8 ACD84362 Human PRO
c 779 14.2 71.0 3265 8 ACD99136 CDNA enco
c 780 14.2 71.0 3265 8 ADA78009 Human sec
c 781 14.2 71.0 3265 8 ACF48878 Human sec
c 782 14.2 71.0 3265 8 ACD09198 Human sec
c 783 14.2 71.0 3265 8 ACF11991 Human sec
c 784 14.2 71.0 3265 8 ACF41225 Human sec
c 785 14.2 71.0 3265 8 ACF15839 Human sec
c 786 14.2 71.0 3265 8 ACF16146 Human sec
c 787 14.2 71.0 3265 8 ADB17126 Human CDN
c 788 14.2 71.0 3265 8 ACD31973 Human sec
c 789 14.2 71.0 3265 8 ACF18781 Human sec
c 790 14.2 71.0 3265 8 ACF09228 Human sec
c 791 14.2 71.0 3265 8 ACF78349 Human sec
c 792 14.2 71.0 3265 8 ACF51948 Human sec
c 793 14.2 71.0 3265 8 ACF26435 Human sec
c 794 14.2 71.0 3265 8 ACF24228 Human sec
c 795 14.2 71.0 3265 8 ACF63539 Human sec
c 796 14.2 71.0 3265 8 ACF50413 Human sec
c 797 14.2 71.0 3265 8 ACH07884 Human sec
c 798 14.2 71.0 3265 8 ACF13690 Human sec
c 799 14.2 71.0 3265 8 ACD41616 Human sec
c 800 14.2 71.0 3265 8 ADA37889 Human CDN

ALIGNMENTS

RESULT 1
AAD21384 strand; DNA; 20 BP.
ID AAD21384;
AC AAD21384;
XX
XX 28-JAN-2002 (first entry)
DT Antisense oligo, HYB 963, directed against human XPA gene.
DE
XX
XX Human; cytotoxin; cancer; transcription coupled repair; TCR;
KW nucleotide existion repair; NER; antisense; cytosaltic;
KW Xeroderma pigmentosum group A; XPA; ss.
XX
XX Homo sapiens.
OS Synthetic.
OS
OS
FH key Location/Qualifiers
FT modified_base 1..20 /*tag= a
FT /mod_base= OTHER

/note= "Phosphorothioate backbone"

FT
XX
XX
PN WO200174346-A2.
XX
PD 11-OCT-2001.
XX
XX 03-APR-2001; 2001WO-US010800.
XX
XX 03-APR-2000; 2000US-0194343P.
XX
XX (HYBR-) HYBRIDON INC.
XX
XX Agrawal S, Kandimalia ER, Bregman DB, Mani S, Lu Y;
XX
XX MPI; 2001-662947/76.
XX
XX
XX Increasing sensitivity of cancer cells to a cytotoxin or oxidizing agent
XX useful for therapy comprises contacting them with oligonucleotides
XX complementary to transcription coupled repair or nucleotide excision
XX repair genes.
XX
XX Claim 12; Page 18; 58pp; English.
XX
XX The present invention relates to a method for potentiating or enhancing
XX the toxic effect of a cytotoxin or oxidizing agent on a cancer cell,
XX comprising contacting the cell with an oligonucleotide complementary to a
XX gene involved in transcription coupled repair (TCR) and nucleotide
XX excision repair (NER) and with a cytotoxin or oxidizing agent. The
XX invention is used to sensitize cancer cells to therapeutic agents. The
XX present sequence is an antisense oligonucleotide directed against
XX Xeroderma pigmentosum group A (XPA) gene
XX
SQ Sequence 20 BP; 4 A; 4 C; 5 G; 7 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
DB 1 GGTCCTACTCATGTTGATG 20

RESULT 2
ABL84566
ID ABL84566 standard; cDNA; 486 BP.
XX
XX ABL84566;
XX
XX 17-MAY-2002 (first entry)
XX
XX Human ovarian cancer related cDNA clone SEQ ID NO:7544.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200192581-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US017756.
XX
XX 26-MAY-2000; 2000US-0207484P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Harlocker SL, Jones R;
XX
XX MPI; 2002-122075/16.
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding

PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
XX Claim 1; SEQ ID NO 7544; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
XX and immunostimulants; and a polypeptide (II) of a ovarian tumour
XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
XX from the 10912 nucleotide sequences as given in ABL77923 to ABL87934,
XX (III) encoding (II) having a sequence (S2), a T cell population of (II),
XX or antigen presenting cells that express (II). (I) has cytostatic
XX activity. An oligonucleotide (IV) that hybridizes to (S1) can be used for
XX detecting ovarian cancer in a patient's biological sample preferably
XX serum or ovarian tissue. The method comprises contacting a biological
XX sample from a patient with (IV), detecting the amount of polynucleotide
XX hybridizing to (IV) and comparing the amount to a predetermined cutoff
XX value and thereby detecting ovarian cancer in the patient, where the
XX amount of polynucleotide hybridizing to (IV) is detected preferably by
XX polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
XX useful for stimulating and/or expanding T cells specific for an ovarian
XX tumour protein comprising contacting T cells with (III) or (II). (III) is
XX useful in design and preparation of ribozyme molecules for inhibiting
XX expression of the tumour polypeptides and proteins in tumour cells; and
XX to isolate a full length gene from a suitable library e.g., a tumour cDNA
XX library using well known techniques

QY 1 GGTCCTACTCATGTTGATG 20
DB 149 GGTCCTACTCATGTTGATG 168

RESULT 3
ABL79139/c
ID ABL79139 standard; cDNA; 506 BP.
XX
XX ABL79139;
XX
XX 17-MAY-2002 (first entry)
XX
XX Human ovarian cancer related cDNA clone SEQ ID NO:2117.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200192581-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US017756.
XX
XX 26-MAY-2000; 2000US-0207484P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Harlocker SL, Jones R;
XX
XX MPI; 2002-122075/16.
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX polypeptide, antibody specific to polypeptide or T cell expressing
XX polypeptide.
XX
XX Claim 1; SEQ ID NO 2117; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers

CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques

SQ Sequence 506 BP; 174 A; 95 C; 104 G; 132 T; 0 U; 1 Other;

Query Match 100.0%; Score 20; DB 6; Length 506;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCCTACTCATGTTGATG 20
|||
Db 246 GGTCCTACTCATGTTGATG 227

RESULT 4
ABL79129
ID ABL79129 standard; cDNA; 513 BP.

XX ABL79129;

XX 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:2107.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX MO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX polypeptide, antibody specific to polypeptide or T cell expressing
XX polypeptide.

XX Claim 1; SEQ ID NO 2107; 489bp; English.

XX The present invention describes a composition (I) comprising: carriers
XX and immunostimulants; and a polypeptide (II) of a ovarian tumour
XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
XX from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
XX (III) encoding (II) having a sequence (S2), a T cell population of (II),
XX or antigen presenting cells that express (II). (I) has cytostatic
XX activity. An oligonucleotide (IV) that hybridises to (S1) can be used for

CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques

SQ Sequence 513 BP; 144 A; 95 C; 100 G; 166 T; 0 U; 8 Other;

Query Match 100.0%; Score 20; DB 6; Length 513;
Best Local Similarity 100.0%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

OY 1 GGTCCTACTCATGTTGATG 20
|||
Db 357 GGTCCTACTCATGTTGATG 376

RESULT 5
AAS63227/C
ID AAS63227 standard; cDNA; 822 BP.

XX AAS63227;

XX 29-JAN-2002 (first entry)

XX Human xeroderma pigmentosum complementation group A (XPA) DNA.

XX DNA mutation-binding protein; nuclease; DNA mismatch; cancer; PCR primer;

XX DNA damage; human xeroderma pigmentosum complementation group; XPF; XPA;

XX XPC; XPE; ERCC4; human MutS homologue 2; hMSH2; MutS; Nuc; MutY; Fpg; ss;

XX Fapy-DNA glycosylase; uracil DNA glycosylase; ung; Tdg; xthA gene; Uvr A;

XX A/G-specific adenine glycosylase; synthetic T4 endonuclease V; T4 endo V;

XX thymine DNA-glycosylase; Uvr B; Uvr C; nth gene; nfo gene; exonuclease;

XX endonuclease.

XX Homo sapiens.

XX MO200173079-A2.

XX 04-OCT-2001.

XX 26-MAR-2001; 2001WO-US009700.

XX 28-MAR-2000; 2000US-0192764P.

XX 29-AUG-2000; 2000US-00650855.

XX (RECG) UNIT CALIFORNIA.

XX Mc Cutchen- Maloney SL;

XX WPI; 2001-656920/75.

XX P-PSDB; AAU69740.

XX Recombinant chimeric protein, useful for detecting and quantifying DNA

XX mutations, e.g. in disease diagnosis, comprises mutation-binding protein

XX and nuclease.

XX Claim 46; Page 62; 128bp; English.

XX Sequences AAS63224-AAS63305 represent cDNA encoding and PCR primers used

XX to amplify cDNA encoding proteins which can be used in the synthesis of

XX chimeric proteins comprising a DNA mutation-binding protein, a linker and

XX a nuclease, by recombinant technology. The chimeric proteins are useful

XX for detection, quantification and mapping of DNA sequence variations

XX including mutations, for example, caused by damage and mismatches. The

CC proteins are able to bind to the site of the DNA mutation and cut it out
CC of the molecule. This is useful for early diagnosis of cancer and other
CC diseases. The proteins used in the invention include human XPF (or
CC ERCC1), human xeroderma pigmentosum complementation groups A, C and B
CC (XPA, XPC and XPD), human Muts homologue 2 (hMSH2), serrata marcescens
CC nuclease (Nuc), Thermus thermophilus Muts, Escherichia coli Fapy-DNA
CC glycosylase (Fpg), uracil DNA glycosylase (ung), A/G-specific adenine
CC glycosylase (Muty), synthetic T4 endonuclease V (T4 endo V), thymine DNA-
CC glycosylase (TDG), E. coli Uvr A, B and C, and E. coli endonucleases and
CC exonucleases
XX
SQ Sequence 822 BP; 280 A; 144 C; 225 G; 173 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTCACACTCATGTTGATG 20
DB 743 GGTCACACTCATGTTGATG 724
RESULT 6
AAD21394/c
ID AAD21394 standard; DNA; 1377 BP.
XX
AC AAD21394;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human Xeroderma pigmentosum group A (XPA) DNA.
XX
KW Human; cytotoxin; cancer; transcription coupled repair; TCR;
KW nucleotide excision repair; NER; cytosolic; chromosome 9;
KW Xeroderma pigmentosum group A; XPA; ds.
XX
OS Homo sapiens.
XX
PN WO200174346-A2.
XX
PD 11-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US010800.
XX
PR 03-APR-2000; 2000US-0194343P.
XX
PA (HYBR-) HYBRIDON INC.
XX
PI Agrawal S, Kandimalia ER, Bregman DB, Mani S, Lu Y;
XX
DR WPI; 2001-662947/76.
XX
PT Increasing sensitivity of cancer cells to a cytotoxin or oxidizing agent
PT useful for therapy comprises contacting them with oligonucleotides
PT complementary to transcription coupled repair or nucleotide excision
PT repair genes.
XX
PS Disclosure; Page 48; 58pp; English.
XX
CC The present invention relates to a method for potentiating or enhancing
CC the toxic effect of a cytotoxin or oxidizing agent on a cancer cell,
CC comprising contacting the cell with an oligonucleotide complementary to a
CC gene involved in transcription coupled repair (TCR) and nucleotide
CC excision repair (NER) and with a cytotoxin or oxidizing agent. The
CC invention is used to sensitize cancer cells to therapeutic agents. The
CC present sequence is a human Xeroderma pigmentosum group A (XPA) gene
CC located on chromosome 9
XX
SQ Sequence 1377 BP; 458 A; 232 C; 358 G; 329 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 4; Length 1377;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCACACTCATGTTGATG 20
DB 769 GGTCACACTCATGTTGATG 750
RESULT 7
AAS34872/c
ID AAS34872 standard; cDNA; 1407 BP.
XX
AC AAS34872;
XX
DT 04-DEC-2001 (first entry)
XX
DE cDNA encoding novel human neoplastic disease associated polypeptide #106.
XX
KW Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX
OS Homo sapiens.
XX
PN WO200155163-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001358.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.

[illegible]

XX	PR	17-NOV-2000;	2000US-0249244P.
XX	PR	17-NOV-2000;	2000US-0249245P.
XX	PR	17-NOV-2000;	2000US-0249264P.
XX	PR	17-NOV-2000;	2000US-0249265P.
XX	PR	17-NOV-2000;	2000US-0249297P.
XX	PR	17-NOV-2000;	2000US-0249299P.
XX	PR	17-NOV-2000;	2000US-0249300P.
XX	PR	01-DEC-2000;	2000US-0250160P.
XX	PR	01-DEC-2000;	2000US-0250391P.
XX	PR	05-DEC-2000;	2000US-0251030P.
XX	PR	05-DEC-2000;	2000US-0251988P.
XX	PR	05-DEC-2000;	2000US-0256719P.
XX	PR	06-DEC-2000;	2000US-0251479P.
XX	PR	08-DEC-2000;	2000US-0251856P.
XX	PR	08-DEC-2000;	2000US-0251868P.
XX	PR	08-DEC-2000;	2000US-0251869P.
XX	PR	08-DEC-2000;	2000US-0251989P.
XX	PR	08-DEC-2000;	2000US-0251990P.
XX	PR	11-DEC-2000;	2000US-0254097P.
XX	PR	05-JAN-2001;	2001US-0259678P.
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Rosen CA, Barash SC, Ruben SM;	
XX	DR	WPI; 2001-465558/50.	
XX	PT	P-PSDB; AAU21673.	
XX	PT	Novel polypeptides and polynucleotides useful as diagnostic reagents to	
XX	PT	diagnose diseases or disorders associated with aberrant expression or	
XX	PT	activity of polypeptides, and for treating cancers, rheumatoid arthritis.	
XX	PS	Claim 4; SEQ ID NO 116; 687bp; English.	
XX	CC	The present invention relates to the isolation of novel human neoplastic	
XX	CC	disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA	
XX	CC	sequences encoding for these polypeptides. The sequences of the invention	
XX	CC	are useful in the diagnosis, treatment, prevention and/or prognosis of	
XX	CC	disorders involving neoplastic disease such as hyperproliferative	
XX	CC	disorders (e.g., leukemia, bone cancer, bladder cancer, brain stem	
XX	CC	glioma, adult liver cancer, childhood cerebellar astrocytoma, or	
XX	CC	Hodgkin's lymphoma). The sequences of the invention may also be useful	
XX	CC	for treating other disorders such as neural disorders, immune system	
XX	CC	disorders, muscular disorders, reproductive disorders, gastrointestinal	
XX	CC	disorders, pulmonary disorders, cardiovascular disorders and renal	
XX	CC	disorders. The polynucleotide sequences of the invention are also useful	
XX	CC	in gene therapy. AAG34767-AAG35050 represent cDNA sequences encoding for	
XX	CC	the novel human neoplastic disease associated polypeptides of the	
XX	CC	invention. Note: The sequence data for this patent did not form part of	
XX	CC	the printed specification, but was obtained in electronic format directly	
XX	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX	SEQ	Sequence 1407 BP; 476 A; 235 C; 365 G; 330 T; 0 U; 1 Other;	
QY	Query Match	100.0%; Score 20; DB 4; Length 1407;	
DJ	Best Local Similarity	100.0%; Pred. No. 4.2;	
DB	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	GATCATTCTCATGTTGATG 20	
DJ			
DB	782	GATCATTCTCATGTTGATG 763	
RESULT 8			
ID	ADCC6030/c		
XX	ADCC6030 standard; cDNA, 1407 BP.		
AC	ADCC6030;		
XX			
DT	18-DEC-2003 (first entry)		
DE	Human neoplastic disease-associated gene 106 cDNA #1.		
XX			

KW Neoplastic disease-associated polypeptide; gene therapy;
KW hyperproliferative disease; cancer; autoimmune disorder; diabetes;
KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;
KW autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;
KW haematologic disorder; anaemia; thrombocytopenia; allergic reaction;
KW asthma; eczema; inflammatory disorder; lechaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; renal disorder;
KW acute glomerulonephritis; end-stage renal disease;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; AIDS; cachexia; anorexia; wound healing;
KW epithelial cell proliferation; Human; ss.
XX
XX Homo sapiens.
XX
XX US2003082758-A1.
XX
XX 01-MAY-2003.
XX
XX 22-MAR-2002; 2002US-00103313.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-020515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-022547P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-022679P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-022668P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0233066P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-023637P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0251989P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764854.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM, Barash SC;
 PI WPI; 2003-786918/74.
 DR P-PSDB; ADC46314.
 XX
 XX New isolated human neoplastic disease-associated polypeptides and
 PT polynucleotides, useful for diagnosing, preventing, prognosticating or
 PT treating medical conditions such as cancer, AIDS, diabetes or Parkinson's
 PT disease.
 PS
 PS Claim 1; SEQ ID NO 116; 302pp; English.
 XX
 CC The invention relates to one of 238 disclosed human neoplastic disease-
 CC associated polypeptides encoded by 171 disclosed cDNA sequences
 CC (including their domains, epitopes, full-length proteins, allelic variants
 CC or species homologues). Also included are there encoding nucleic acids, a
 CC recombinant vector comprising the nucleic acid, a recombinant host cell
 CC comprising the nucleic acid (expressing the protein), an isolated
 CC antibody that binds specifically to the isolated polypeptide, preventing,
 CC treating or ameliorating a medical condition, diagnosing a pathological
 CC condition or a susceptibility to a pathological condition in a subject,
 CC identifying a binding partner to the polypeptide, identifying an activator
 CC in a biological assay, and the gene corresponding to the cDNA sequence.
 CC The polypeptides, polynucleotides and antibodies are useful for
 CC detecting, preventing, diagnosing, prognosticating, treating or
 CC ameliorating medical conditions such as hyperproliferative diseases or
 CC cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis,
 CC systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis
 CC or haemolytic anaemia), haematopoietic or haematologic disorders (e.g.
 CC anaemia or thrombocytopenia), allergic reactions including asthma or
 CC eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,
 CC
 Query Match 100.0%; Score 20; DB 9; Length 1407;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTTGATG 20
 DB 782 GGTCCATCTCATGTTGATG 763
 RESULT 9
 AB235476
 ID AB235476 standard; cDNA; 4670 BP.
 XX
 AC AB235476;
 XX
 DT 05-FEB-2003 (first entry)
 XX
 DE Human gene expression profile polynucleotide SEQ ID NO 587.
 XX
 KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KW gene expression; gene; ss.
 XX
 OS Homo sapiens.
 XX

PN WO200274979-A2.
 XX
 PD 26-SEP-2002.
 XX
 PP 20-MAR-2002; 2002MO-US008456.
 XX
 PR 20-MAR-2001; 2001US-0276947P.
 XX
 PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 XX
 PI Wan J, Wang Y;
 DR WPI; 2002-740862/80.
 XX
 XX New gene expression profile generated from primary, endothelial,
 PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer.
 XX
 PS Disclosure; Page 716-718; 850pp; English.
 XX
 CC The invention relates to a gene expression profile comprising one or more
 CC genes (AB234889-AB235692) and generated from a cell type. The cell type
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
 CC endothelium, myometrium microvascular endothelium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 CC for determining the level of RNA expression for a sample, determining the
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 CC expression profile is useful in identifying disease pathologies involving
 CC alterations of gene expression. The assessment of expression profiles may
 CC provide meaningful information with respect to tumour type and stage,
 CC treatment methods, and prognosis. The gene or protein expression profile
 CC may also be used for creating microarrays. The microarray is useful for
 CC genetic and physical mapping of genomes. DNA sequencing, genetic or
 CC medical diagnosis, genotyping of organisms, confirming cell or tissue
 CC identifications and in identifying promising antibiotics, antiviral or
 CC antifungal agents
 SQ Sequence 4670 BP; 1072 A; 1094 C; 977 G; 1055 T; 0 U; 472 Other;
 Query Match 100.0%; Score 20; DB 6; Length 4670;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTTGATG 20
 DB 695 GGTCCATCTCATGTTGATG 714
 RESULT 10
 ABX05799
 ID ABX05799 standard; DNA; 1029 BP.
 XX
 AC ABX05799;
 XX
 DT 27-OCT-2003 (revised)
 XX
 DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain coding region #87.
 XX
 KW Gene; ds; bacterial meningitis; pneumonia; sepsis; otitis media;
 KW ear infection; antiinflammatory; antibacterial; immunostimulant;
 KW auditory; respiratory; gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 XX

XX	MO200277021-A2.
XX	03-OCT-2002.
PF	27-MAR-2002; 2002MO-IB002163.
PR	27-MAR-2001; 2001GB-00007658.
XX	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
XX	Maignani V, Tettelin H, Fraser C;
XX	WPL, 2003-040579/03.
DR	P-PSDB; AB000520.
XX	New proteins and nucleic acid molecules from <i>Streptococcus pneumoniae</i> , PT useful as medicaments for treating or preventing a disease or infection PT due to <i>Streptococcus bacteria</i> , such as pneumonia, sepsis, otitis media or PT ear infection.
PS	Claim 6; SEQ ID NO 173; 56pp; English.
XX	The invention relates to a protein comprising or having at least 50%
CC	identity to any of the 2469 amino acid sequences, identified in the
CC	specification (available on a computer readable format), or its fragment,
CC	expressed from 2469 of 2489 identified DNA coding regions from the
CC	<i>Streptococcus pneumoniae</i> type 4 strain genomic sequence appearing as
CC	AB556454. Also included are an antibody which binds one of the proteins,
CC	treating a patient by administering the protein, DNA or antibody (in a
CC	composition), a kit comprising first and second primers, which are the
CC	nucleic acid cited above or fragments between nucleotides 8-100 of a
CC	sequence not defined in the specification, for amplifying a target
CC	sequence contained within a <i>Streptococcus</i> nucleic acid sequence, where
CC	the first primer is substantially complementary to the target sequence
CC	and the second primer is substantially complementary to the complement of
CC	the target sequence, and where the parts of the primers having
CC	substantial complementarity define the termini of the target sequence to
CC	be amplified, assay comprising contacting a test compound with the
CC	protein, and determining whether the test compound binds to the protein
CC	and a <i>Streptococcus pneumoniae</i> bacterium, where one or more genes
CC	encoding the proteins has been rendered inactive. The proteins, nucleic
CC	acid molecules, antibody and compositions are useful as medicaments for
CC	treating or preventing a disease or infection due to <i>Streptococcus</i>
CC	bacteria, particularly <i>S. pneumoniae</i> , such as pneumonia, sepsis, otitis
CC	media or ear infection. They are also useful in developing vaccines,
CC	diagnostics and antibiotics. The methods are useful for identifying
CC	immunodominant proteins. The present sequence is one of the 2489
CC	identified coding region from the genomic sequence. Note: The sequence
CC	data for this patent did not form part of the printed specification, but
CC	was obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
CC	standardise OS field)
XX	Sequence 1029 BP; 305 A; 221 C; 212 G; 291 T; 0 U; 0 Other;
XX	Query Match 84.0%; Score 16.8; DB 7; Length 1029;
XX	Best Local Similarity 90.0%; Pred. No. 1.6e+02;
QY	Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0
DB	1 GGTCCATACCTCATGTGATG 20
	975 GGTCCATGCTCATGTGATG 994
RESULT 11	
ID	AAVS2173/c
XX	AAVS2173 standard; DNA; 14273 BP.
AC	AAVS2173;
XX	
XT	23-OCT-1998 (first entry)
XX	

Accession	Fragment Name	Begin	End	LOCUS	ABSS6454	Accession	ABSS6454
ABSS6454_00	Fragment Name	1	110000				
ABSS6454_01		100001	210000				
ABSS6454_02		200001	310000				
ABSS6454_03		300001	410000				
ABSS6454_04		400001	510000				
ABSS6454_05		500001	610000				
ABSS6454_06		600001	710000				

WP ABS56454_07 700001 810000
WP ABS56454_08 800001 910000
WP ABS56454_09 900001 1010000
WP ABS56454_10 1000001 1110000
WP ABS56454_11 1100001 1210000
WP ABS56454_12 1200001 1310000
WP ABS56454_13 1300001 1410000
WP ABS56454_14 1400001 1510000
WP ABS56454_15 1500001 1610000
WP ABS56454_16 1600001 1710000
WP ABS56454_17 1700001 1810000
WP ABS56454_18 1800001 1910000
WP ABS56454_19 1900001 2010000
WP ABS56454_20 2000001 2110000
WP ABS56454_21 2100001 2162598
ID ABS56454 Standard; DNA; 2162598 BP.
XX
XX ABS56454;
AC
XX
XX 27-OCT-2003 (revised)
DT
XX 10-FEB-2003 (first entry)
DT
XX
XX Streptococcus pneumoniae type 4 strain complete genome.
DE
XX
XX da; bacterial meningitis; pneumonia; sepsis; otitis media; genome;
KW ear infection; antiinflammatory; antibacterial; immunostimulant;
KW auditory; respiratory; gene therapy; vaccine.
XX
XX OS Streptococcus pneumoniae; type 4 strain.
XX
XX PN WO200277021-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 27-MAR-2002; 2002WO-IB002163.
XX
XX PR 27-MAR-2001; 2001GB-00007658.
XX
XX PA (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX
XX PI Masignani V, Tettelin H, Fraser C;
XX
XX DR MPI; 2003-040579/03.
XX
XX PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX
XX PS Claim 17; SEQ ID NO 4979; 56bp; English.
XX
XX CC The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC ABS36454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the proteins has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to streptococcus

CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is the Streptococcus
CC pneumoniae type 4 strain genome sequence. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIP0 at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
XX SQ Sequence 2162598 BP; 654373A; 427176C; 431369G; 649680T; 0U; 00cher;
SQ
Query Match 84.0%; Score 16.8; DB 7; Length 110000;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGTCATGCTCATGTGATG 20
DB 67573 GGTCATGCTCATGTGATG 67554
RESULT 13
ID ABT13966/c
XX ABT13966 standard; DNA; 6765 BP.
XX
XX AC ABT13966;
XX
XX DT 13-FEB-2003 (first entry)
XX
XX DE Human Cryptovirus strain BBR DNA region SEQ ID No 23.
XX
XX KW Neuroprotective; virucide; anticonvulsant; antiinflammatory; Cryptovirus;
KW idiopathic; neurological; neurodegenerative; neuropsychological; vaccine;
KW epilepsy; neuropsychiatric disorder; subacute sclerosing panencephalitis;
KW epileptiform disease; multiple sclerosis; chronic fatigue syndrome;
KW primary lymphadenopathy-associated illness; gene therapy; de.
XX
XX OS Rubulavirus sp.
XX
XX PN WO200277211-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 07-FEB-2002; 2002WO-US004117.
XX
XX PR 07-FEB-2001; 2001US-0267253P.
XX
XX PA (CRYP-) CRYPTIC AFFILICTIONS LLC.
XX
XX PI Robbins SJ;
XX
XX DR MPI; 2003-040586/03.
XX P-PSDB; ABO18512.
XX
XX PT New nucleic acid useful for diagnosing and treating idiopathic
PT neurological disorders, including epileptiform diseases, e.g. epilepsy,
PT and lymphadenopathy-associated illnesses, and in screening of potential
PT new antiviral drugs.
XX
XX XX Disclosure; Page 225-235; 262pp; English.
XX
XX CC The invention relates to an isolated nucleic acid comprising: contiguous
CC nucleotide positions 1-15246 of a sequence comprising 15246 bp fully
CC defined in the specification; a nucleotide sequence complementary to the
CC 15246bp sequence; or a Cryptovirus-specific fragment of the 15246bp or
CC its complement, at least about 5 nucleotides long. The nucleic acid is
CC useful in diagnosing and treating many idiopathic neurological,
CC neurodegenerative, neuropsychological and neuropsychiatric disorders,
CC including epileptiform diseases (e.g. epilepsy, multiple sclerosis,
CC chronic fatigue syndrome or subacute sclerosing panencephalitis) and
CC primary lymphadenopathy-associated illnesses, and in research and
CC development, including screening of potential new antiviral drugs. The
CC nucleic acid, protein and the viral particle are useful in manufacturing

CC a vaccine. The protein is also used in producing a Cryptovirus-specific
CC antibody. The antibody may also be used in manufacturing a medicament for
CC the treatment of Cryptovirus infections. The polynucleotides of the
CC invention can be used to treat disorders by gene therapy. This
CC polynucleotide sequence represents a DNA region of the human Cryptovirus
CC strain BBR of the invention
SQ Sequence 6765 BP; 2139 A; 1351 C; 1357 G; 1918 T; 0 U; 0 Other;
Query Match 82.0%; Score 16.4; DB 7; Length 6765;
Best Local Similarity 94.4%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCCATACCTCATGTTGATG 20
DB 4712 TCCATACCTCATGTTGATG 4695
RESULT 14
ID ABR13954/C
ABR13954 standard; DNA; 15246 BP.
XX
AC ABR13954;
XX
DT 13-FEB-2003 (first entry)
XX
DE Human Cryptovirus strain BBR coding sequence SEQ ID NO 1.
XX
KW Neuroprotective; virucide; anticonvulsant; antiinflammatory; Cryptovirus;
KW idiopathic; neurodegenerative; neuropsychological; vaccine;
KW epilepsy; neuropsychiatric disorder; subacute sclerosing panencephalitis;
KW epileptiform disease; multiple sclerosis; chronic fatigue syndrome;
KW primary lymphadenopathy-associated illness; gene therapy; gene; ds.
OS Rubulavirus sp.
XX
PN WO20027211-A2.
XX
PD 03-OCT-2002.
XX
PF 07-FEB-2002; 2002WO-US004117.
XX
PR 07-FEB-2001; 2001US-0267253P.
PA (CRYP-) CRYPTIC AFFLICTIONS LLC.
PI Robbins SJ;
XX
DR MPI; 2003-040586/03.
XX
PT New nucleic acid useful for diagnosing and treating idiopathic
PT neurological disorders, including epileptiform diseases, e.g. epilepsy,
PT and lymphadenopathy-associated illnesses, and in screening of potential
PT new antiviral drugs.
XX
PS Claim 1, Fig 9; 262pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising: contiguous
CC nucleotide positions 1-15246 of a sequence comprising 15246 bp fully
CC defined in the specification; a nucleotide sequence complementary to the
CC 15246bp sequence; or a Cryptovirus-specific fragment of the 15246bp or
CC its complement, at least about 5 nucleotides long. The nucleic acid is
CC useful in diagnosing and treating many idiopathic neurological,
CC neurodegenerative, neuropsychological and neuropsychiatric disorders,
CC including epileptiform diseases (e.g. epilepsy, multiple sclerosis,
CC chronic fatigue syndrome or subacute sclerosing panencephalitis) and
CC primary lymphadenopathy-associated illnesses, and in research and
CC development, including screening of potential new antiviral drugs. The
CC nucleic acid, protein and the viral particle are useful in manufacturing
CC a vaccine. The protein is also used in producing a Cryptovirus-specific
CC antibody. The antibody may also be used in manufacturing a medicament for
CC the treatment of Cryptovirus infections. The polynucleotides of the
CC invention can be used to treat disorders by gene therapy. This

CC polynucleotide sequence represents the coding DNA for the human
CC Cryptovirus strain BBR protein of the invention
XX
SQ Sequence 15246 BP; 4753 A; 3273 C; 3131 G; 4089 T; 0 U; 0 Other;
Query Match 82.0%; Score 16.4; DB 7; Length 15246;
Best Local Similarity 94.4%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCCATACCTCATGTTGATG 20
DB 13125 TCCATACCTCATGTTGATG 13108
RESULT 15
ABA90521_15
Continuation (16 of 24) of ABA90521 from base 1500001 (Genomic sequence of Lactococcus 1
WP Sequence split into 24 fragments LOCUS ABA90521 Accession ABA90521
WP Fragment Name Begin End
WP ABA90521_00 1 110000
WP ABA90521_01 100001 210000
WP ABA90521_02 200001 310000
WP ABA90521_03 300001 410000
WP ABA90521_04 400001 510000
WP ABA90521_05 500001 610000
WP ABA90521_06 600001 710000
WP ABA90521_07 700001 810000
WP ABA90521_08 800001 910000
WP ABA90521_09 900001 1010000
WP ABA90521_10 1000001 1110000
WP ABA90521_11 1100001 1210000
WP ABA90521_12 1200001 1310000
WP ABA90521_13 1300001 1410000
WP ABA90521_14 1400001 1510000
WP ABA90521_15 1500001 1610000
WP ABA90521_16 1600001 1710000
WP ABA90521_17 1700001 1810000
WP ABA90521_18 1800001 1910000
WP ABA90521_19 1900001 2010000
WP ABA90521_20 2000001 2110000
WP ABA90521_21 2100001 2210000
WP ABA90521_22 2200001 2310000
WP ABA90521_23 2300001 2365589
Query Match 82.0%; Score 16.4; DB 6; Length 110000;
Best Local Similarity 94.4%; Pred. No. 4.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCCATACCTCATGTTGATG 20
DB 51686 TCCATACCTCATGTTGATG 51703
RESULT 16
ABV97710/C
ID ABV97710 standard; cDNA; 431 BP.
XX
AC ABV97710;
XX
DT 14-JAN-2003 (first entry)
XX
DE Human pancreatic cancer expressed cDNA SEQ ID NO 3118.
XX
KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW cytostatic; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200260317-A2.
XX
PD 08-AUG-2002.
XX
PF 30-JAN-2002; 2002WO-US002781.

```

XX 30-JAN-2001; 2001US-0265305P.
PR 31-JAN-2001; 2001US-0265682P.
PR 09-FEB-2001; 2001US-0267568P.
PR 21-MAR-2001; 2001US-0278651P.
PR 28-APR-2001; 2001US-0287112P.
PR 16-MAY-2001; 2001US-0291631P.
PR 12-JUL-2001; 2001US-0305484P.
PR 20-AUG-2001; 2001US-0313999P.
PR 27-NOV-2001; 2001US-0333626P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Heppler WT, Jiang Y;
XX
XX WPI; 2002-627435/67.
XX
XX DR New isolated polynucleotide and pancreatic tumor polypeptides, useful for
XX PT diagnosing, preventing and/or treating cancer, particularly pancreatic
XX PT cancer.
XX
XX PS Claim 1; SEQ ID NO 3118; 300bp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated polynucleotide (1) comprising: (a)
XX CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
XX CC complements of (a); (c) sequences consisting of at least 20 contiguous
XX CC residues of (a); (d) sequences that hybridize to (a), under moderately
XX CC stringent conditions; (e) sequences having at least 75% or 90% identity
XX CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
XX CC ABP68633) encoded by (1) and oligonucleotide can be used to detect cancer
XX CC in a patient and compositions comprising polypeptides, polynucleotides,
XX CC antibodies, fusion proteins, T cell populations and antigen presenting
XX CC cells expressing the polypeptide are useful in treating pancreatic cancer
XX CC and stimulating an immune response. The polynucleotides can be used as
XX CC probes or primers for nucleic acid hybridisation, in the design and
XX CC preparation of ribozyme molecules for inhibiting expression of the tumour
XX CC polypeptides and proteins in the tumour cells, in vaccines and for gene
XX CC therapy. Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WFO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 431 BP; 150 A; 63 C; 60 G; 156 T; 0 U; 2 Other;
XX
XX Query Match 79.0%; Score 15.8; DB 6; Length 431;
XX Best Local Similarity 89.5%; Pred. No. 4.6e+02;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 2 GTCCATCTCATGTGATG 20
XX DB 194 GTCCATATCATCAAGTGTGATG 176
XX
XX RESULT 17
XX ACH25004/c
XX ID ACH25004 standard; cDNA; 462 BP.
XX
XX AC ACH25004;
XX
XX DT 13-OCT-2003 (first entry)
XX
XX DE Human adult ovary cDNA #3384.
XX
XX KW Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;
XX KW genome mapping; biodiversity; genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN US2003073623-A1.
XX
XX PD 17-APR-2003.
XX
XX PF 30-JUL-2001; 2001US-00918995.
XX

```

```

PR 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
XX DR New polynucleotide sequences obtained from various cDNA libraries, useful
XX PT as hybridization probes, as oligomers for PCR, for chromosome and gene
XX PT mapping, in the recombinant production of protein, or in generating
XX PT antisense DNA or RNA.
XX
XX PS Claim 1; SEQ ID NO 12216; 44bp; English.
XX
XX CC The invention relates to an isolated polynucleotide comprising any one of
XX CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX CC determined by the technique of SHH (sequencing by hybridisation). Also
XX CC included is a purified polypeptide comprising a sequence corresponding to
XX CC a reading frame of the novel polynucleotide. The nucleic acid sequences
XX CC are useful in diagnostics as expressed sequence tags (EST) for
XX CC identifying expressed genes or for physical mapping of the human genome,
XX CC in forensics, in assessing biodiversity, or in identifying mutations
XX CC responsible for genetic disorders and other traits. The nucleotide
XX CC sequences are also useful as hybridisation probes, as oligomers for PCR,
XX CC for chromosome and gene mapping, in the recombinant production of
XX CC protein, or in generating antisense DNA or RNA. The purified polypeptide
XX CC is useful for generating antibodies specific for it. The present sequence
XX CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX CC for this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html?docID=20030073623
XX
XX SQ Sequence 462 BP; 104 A; 120 C; 124 G; 112 T; 0 U; 2 Other;
XX
XX Query Match 79.0%; Score 15.8; DB 8; Length 462;
XX Best Local Similarity 89.5%; Pred. No. 4.7e+02;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 2 GTCCATCTCATGTGATG 20
XX DB 446 GTCCCTTCATGTTGATG 428
XX
XX RESULT 18
XX AAA16057
XX ID AAA16057 standard; DNA; 614 BP.
XX
XX AC AAA16057;
XX
XX DT 14-JUN-2000 (first entry)
XX
XX DE Human colon cancer differentially expressed nucleotide sequence #62.
XX
XX KW Colon cancer; detect; differential expression; human; treatment;
XX KW detect mutation; non-invasive diagnostic method; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200012702-A2.
XX
XX PD 09-MAR-2000.
XX
XX PF 30-AUG-1999; 99WO-US019424.
XX
XX PR 31-AUG-1998; 98US-0098639P.
XX PR 27-JAN-1999; 99US-0117393P.
XX
XX (FARB ) BAYER CORP.
XX

```


XX Endege WO, Steinmann KE, Astle JH, Burgess CC, Carroll E;
PI Carino TJ, Divedi P, Ford DM, Lewis ME, Molino GA, Monahan JE;
PI Schlegel R;
XX WPI; 2000-256641/22.
XX Novel nucleic acids and proteins for identifying therapeutic agents
XX useful for treating and diagnosing cancer, especially colon cancer.
PS Claim 16; Page 149; 345pp; English.
XX This sequence represents a human nucleotide sequence which is
CC differentially expressed in colon cancer cells compared to the expression
CC levels in normal cells. The nucleotide sequence can be used as a source
CC of primers and probes. The nucleotide sequence is useful for determining
CC the phenotype of a cell by detecting the differential expression of the
CC sequence relative to a normal cell. The probes derived from the sequence
CC can also be used to determine the phenotype of cells in a sample. Probes
CC and antibodies which hybridize to the nucleotide sequence can also be
CC used to determine the phenotype of a cell. The primers are useful for
CC detecting a mutation in a test nucleotide sequence and also for detecting
CC cancer, preferably colon cancer. Antibodies against the protein encoded
CC by the nucleotide sequence can also be used in a method to detect colon
CC cancer. The diagnostic method is non-invasive and accurate for diagnosing
CC colon cancer at an early stage
SQ Sequence 614 BP; 202 A; 91 C; 99 G; 214 T; 0 U; 8 Other;
OY Query Match 79.0%; Score 15.8; DB 3; Length 614;
Best Local Similarity 89.5%; Pred. No. 4.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 2 GTCCATCTCATGTTGATG 20
466 GTCCATATCAAGTGTATG 484
RESULT 19
ABL23928
ID ABL23928 standard; DNA; 2830 BP.
XX ABL23928;
XX AC
XX 26-MAR-2002 (first entry)
XX DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 23257.
XX DE
XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW
XX pharmaceutical; gene; de.
XX PA
XX Drosophila melanogaster.
XX OS
XX WO200171042-A2.
XX PN
XX 27-SEP-2001.
XX PD
XX 23-MAR-2001; 2001WO-US009231.
XX PF
XX 23-MAR-2000; 2000US-0191637P.
XX PR
XX 11-JUL-2000; 2000US-00614150.
XX RA
XX (PEKE) PE CORP NY.
XX PI
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX Claim 1; SEQ ID NO 23257; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-
CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 2830 BP; 740 A; 666 C; 674 G; 750 T; 0 U; 0 Other;
OY Query Match 79.0%; Score 15.8; DB 4; Length 2830;
Best Local Similarity 89.5%; Pred. No. 6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 1 GATCCATCTCATGTTGAT 19
2480 GGTTCATAGTCATGTTGAT 2498
RESULT 20
AB11966/C
ID AB11966 standard; CDNA; 6454 BP.
XX AB11966;
XX AC
XX 26-MAR-2002 (first entry)
XX DT
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 30380.
XX DE
XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW
XX pharmaceutical; gene; ss.
XX OS
XX Drosophila melanogaster.
XX PN
XX WO200171042-A2.
XX PD
XX 27-SEP-2001.
XX PF
XX 23-MAR-2001; 2001WO-US009231.
XX PR
XX 23-MAR-2000; 2000US-0191637P.
XX PR
XX 11-JUL-2000; 2000US-00614150.
XX PA
XX (PEKE) PE CORP NY.
XX PI
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR
XX P-PsDB; ABB67863.
XX PT
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX Claim 1; SEQ ID NO 30380; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-
CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 6454 BP; 1917 A; 1358 C; 1277 G; 1902 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 6454;
 Best Local Similarity 89.5%; Pred. No. 6.7e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
 |||||
 DB 1006 GTTCATAGTCATGATGATG 988

RESULT 21

ABLO3787/c
 ID ABL03787 standard; cDNA; 8114 BP.

AC ABL03787;
 XX

DT 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5843.
 XX

KW Drosophila; developmental biology; cell signalling; insecticide;
 XX

KW pharmaceutical; gene; ss.
 XX

OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX

PD 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US009231.
 XX

PR 23-MAR-2000; 2000US-0191637P.
 XX

PR 11-JUL-2000; 2000US-00614150.
 XX

PA (PEKE) PE CORP NY.
 XX

PI Venter JC, Adams M, Li PWD, Myers EW;
 XX

DR WPI; 2001-656860/75.
 XX

P-PSDB; ABB59684.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions.
 XX

PS Claim 1; SEQ ID NO 5843; 21bp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SO Sequence 8114 BP; 2159 A; 2177 C; 2263 G; 1515 T; 0 U; 0 Other;
 XX

Query Match 79.0%; Score 15.8; DB 4; Length 8114;
 Best Local Similarity 89.5%; Pred. No. 6.9e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGAT 19
 |||||
 DB 7755 GGTCATAGTCATGTTGAT 7737

RESULT 22

ABLO3786
 ID ABL03786 standard; cDNA; 18737 BP.
 XX

AC ABL03786;
 XX

XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 5840.
 XX
 DE Drosophila; developmental biology; cell signalling; insecticide;
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 DR P-PSDB; ABB59683.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions.
 XX
 PS Claim 1; SEQ ID NO 5840; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SO Sequence 18737 BP; 4552 A; 4551 C; 4677 G; 4957 T; 0 U; 0 Other;
 XX

Query Match 79.0%; Score 15.8; DB 4; Length 18737;
 Best Local Similarity 89.5%; Pred. No. 7.7e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGAT 19
 |||||
 DB 1360 GTTCATAGTCATGTTGAT 1378

RESULT 23

ABLO396/c
 ID ABL0396 standard; cDNA; 24401 BP.
 XX

AC ABL0396;
 XX

DT 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 4670.
 XX

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX

OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX

PD 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PMD, Myers EM;
XX WPI: 2001-656860/75.
XX P-PSDB; ABB59293.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX
XX Claim 1; SEQ ID NO 4670; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB16511), expressed DNA
CC sequences (AB161840-AB16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 24401 BP; 7446 A; 4747 C; 4751 G; 7457 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 4; Length 24401;
Best Local Similarity 89.5%; Pred. No. 8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 2 GTCCATCTCATGTGATG 20
Db 12878 GTCAATCTCATGTGATG 12860
RESULT 24
ABK42099
ID ABK42099 standard; cDNA; 551 BP.
XX
AC ABK42099;
XX
DT 21-MAY-2002 (first entry)
XX
DE cDNA encoding novel human connective tissue related polypeptide #487.
XX
XX Human; connective tissue related disorder; cancer; gene therapy;
KW cytosolic; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200155343-A1.
XX
XX
XX 02-AUG-2001.
XX
XX
XX 17-JAN-2001; 2001WO-US001322.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-020515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX

PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 14-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
XX

PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-565190/63.
DR P-PSDB; AAU86921.
XX
XX Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder such
PT as cancer or rheumatoid arthritis.
XX
XX
PS Claim 4; SEQ ID NO 497; 673bp; English.
XX
XX The present invention relates to the isolation of novel human connective
CC tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
CC (cDNA and genomic) sequences encoding them. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of diseases associated with connective tissue(s), including
CC cancer. The polynucleotide sequences of the invention are also useful in
CC gene therapy. ABK41613-ABK42101 represent cDNA sequences encoding the
CC novel human connective tissue related polypeptides. Note: The sequence

CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 551 BP; 143 A; 114 G; 194 T; 0 U; 6 Other;
Query Match 77.0%; Score 15.4; DB 4; Length 551;
Best Local Similarity 94.1%; Pred. No. 7.66+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 3 TCCATCTCATGTTGAT 19
|||||
Db 261 TCCATCTCATGTTGAT 277
|||||
RESULT 25
ADB59766 standard; cDNA, 551 BP.
XX
XX ADB59766;
AC
XX
DT 04-DEC-2003 (first entry)
XX
XX
DE Connective tissue related polynucleotide #487.
XX
XX cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
XX antiarteriosclerotic; immunosuppressive; antirheumatic; antiarrhythmic;
XX antiinflammatory; antiallergic; antiasthmatic; dermatological;
XX nephropic; virucide; fungicide; antibacterial; antiparasitic;
XX gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
XX systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
XX cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;
XX Alzheimer's disease; Parkinson's disease; cardiovascular disease;
XX atherosclerosis; myocarditis; cardiopulmonary bypass complication;
XX autoimmune disease; multiple sclerosis; allergic reaction; asthma;
XX rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
XX gastrointestinal disorder; inflammatory bowel disease;
XX organ transplant rejection; immune system disorder; Bruton's disease;
XX X-linked lymphoproliferative syndrome;
XX B-cell lymphoproliferative disorder; HIV; AIDS; infection;
XX chromosome identification; chromosome mapping;
XX connective tissue related polynucleotide; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX US2003054375-A1.
PN
XX
XX 20-MAR-2003.
PD
XX
XX 07-MAR-2002; 2002US-00092154.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.

```
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227709P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231422P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235634P.
PR 27-SEP-2000; 2000US-0235636P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241877P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0255719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764847.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI
XX
XX MPI; 2003-634869/60.
DR P-PSDB; ADB60255.
XX
XX The invention describes an isolated nucleic acid molecule (1), which
XX comprises a sequence that is at least 95 % identical to a connective
XX tissue-related polynucleotide encoding connective tissue antigens (CTA).
XX The polypeptide or polynucleotide is useful for preventing, treating, or
XX ameliorating medical conditions in a mammal. The connective tissue
XX polypeptides, polynucleotides and antibodies are particularly useful for
XX treating, preventing and/or prognosing disorders of connective tissues
XX (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,
XX scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or
XX neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.
XX Alzheimer's disease, or Parkinson's disease), cardiovascular diseases
XX (e.g. atherosclerosis, myocarditis or cardiomyopathy) by-pass
XX complications), autoimmune diseases (e.g. systemic lupus erythematosus,
XX rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.
XX
```

Query Match

77.0%; Score 15.4; DB 8; Length 551;

Best Local Similarity 94.1%; Pred. No. 7.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGAT 19
| | | | | | | | | | | | | | | | | | | | | |
DB 261 TCCATTCATCATGTTGAT 277

RESULT 26

ABV55825/c
ID ABV55825 standard; cDNA; 609 BP.

AC ABV55825;

DT 17-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 55816.

KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

PR 17-FEB-2000; 2000US-018319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JR;

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 10778; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (c) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 609 BP; 211 A; 104 C; 87 G; 204 T; 0 U; 3 Other;

Query Match 77.0%; Score 15.4; DB 5; Length 609;

Best Local Similarity 94.1%; Pred. No. 7.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGAT 19
| | | | | | | | | | | | | | | | | | | | | |
DB 468 TCCATACATCATGTTGAT 452

RESULT 27

AAH52789
ID AAH52789 standard; DNA; 633 BP.

AC AAH52789;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:971.

KM Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;

KW endocarditis; ds.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US030782.

PR 09-NOV-1999; 99US-0164258P.

PA (GLAXO) GLAXO GROUP LTD.

PI Kimmeryl WJ;

DR WPI; 2001-316495/33.

DR P-PsDB; AAH81939.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

PT useful for vaccinating against infections, e.g. endocarditis.

PS Claim 8; Page 285; 2188pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

CC (II), given in AAH81454 to AAH83120, from Staphylococcus epidermidis. (I)

CC and (II) can have antibacterial activity and therefore can be used in

CC vaccination. The nucleic acids (I) may be used to produce the S.

CC epidermidis polypeptides (II) via the production of vectors containing

CC them which are used to produce hosts cells which express the

CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be

CC used to vaccinate subjects and to raise antibodies against the bacteria.

CC The polypeptides may also be used to assay for other inhibitors of their

CC activity and therefore identify compounds that may be used for the

CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to

CC AAH5090 represent specifically claimed S. epidermidis genomic DNA

CC polynucleotide sequences from the present invention. AAH5091 to AAH5098

CC represent oligonucleotide sequences and primers which are used in the

CC exemplification of the present invention. N.B. The present invention

CC specifically claims all the polynucleotide sequences given in the

CC sequence listing of the present specification, however the sequence

CC listing only goes up to SEQ ID NO:4454 so even though sequences are given

CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present

CC for SEQ ID NO:4455 to 4464

Sequence 633 BP; 238 A; 87 C; 124 G; 184 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 633;

Best Local Similarity 94.1%; Pred. No. 7.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGA 18
| | | | | | | | | | | | | | | | | | | | | |
DB 385 GTCCATACATCATGTTGA 401

RESULT 28

AAH53645
ID AAH53645 standard; DNA; 675 BP.

AC AAH53645;

DT 03-SEP-2001 (first entry)

XX S. epidermidis open reading frame nucleotide sequence SEQ ID NO:2683.
DE Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
XX endocarditis; ds.
KW
XX Staphylococcus epidermidis.
OS
XX WO200134809-A2.
PN
XX 17-MAY-2001.
PD
XX 09-NOV-2000; 2000WO-US030782.
PF
XX 09-NOV-1999; 99US-0164258P.
PR
XX (GLAX) GLAXO GROUP LTD.
PA
XX Kimerly WJ;
PI
XX MPI: 2001-316495/33.
DR
XX P-PSDB; AAG82795.
PT
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
XX
XX Claim 8; Page 702; 2188pp; English.
XX
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG8120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5509 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
XX
SQ Sequence 675 BP; 260 A; 91 C; 132 G; 192 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 675;
Best Local Similarity 94.1%; Pred. No. 7.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCGATCACTCATGTTGA 18
DB 427 GTCGATCACTCATGTTGA 443

RESULT 29
ABN90925
ID ABN90925 standard; DNA; 675 BP.
XX
XX ABN90925;
AC
XX
XX 24-JUL-2002 (first entry)
DT
XX
XX Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:388.
DE
XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX

OS Staphylococcus epidermidis.
XX
XX US6380370-B1.
PN
XX 30-APR-2002.
PD
XX
XX 13-AUG-1998; 98US-00134001.
PF
XX 14-AUG-1997; 97US-0055779P.
PR
XX 08-NOV-1997; 97US-0064964P.
PA
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
PI
XX MPI: 2002-381255/41.
DR
XX P-PSDB; ABP38380.
DR
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT polypeptide, useful for diagnosing and treating bacterial infections.
PT
XX
XX Disclosure; SEQ ID NO 388; 267pp; English.
PS
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences can
CC also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the USPTO web site
XX
SQ Sequence 675 BP; 260 A; 90 C; 132 G; 193 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 6; Length 675;
Best Local Similarity 94.1%; Pred. No. 7.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCGATCACTCATGTTGA 18
DB 427 GTCGATCACTCATGTTGA 443

RESULT 30
AAC46198/C
ID AAC46198 standard; DNA; 2132 BP.
XX
XX AAC46198;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 49255.
DE
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 98US-0121825P.
PR
XX 05-MAR-1999; 98US-0123180P.
PR
XX 09-MAR-1999; 99US-0125788P.
PR
XX 23-MAR-1999; 99US-0125788P.
PR
XX 25-MAR-1999; 99US-0126264P.
PR
XX 29-MAR-1999; 99US-0126785P.
PR

PR 01-APR-1999;	99US-0127462P.	PR 19-JUL-1999;	99US-0144331P.
PR 06-APR-1999;	99US-0128234P.	PR 19-JUL-1999;	99US-0144333P.
PR 08-APR-1999;	99US-0128714P.	PR 19-JUL-1999;	99US-0144337P.
PR 16-APR-1999;	99US-0129845P.	PR 19-JUL-1999;	99US-0144339P.
PR 19-APR-1999;	99US-0130077P.	PR 19-JUL-1999;	99US-0144335P.
PR 21-APR-1999;	99US-0130449P.	PR 20-JUL-1999;	99US-0144352P.
PR 23-APR-1999;	99US-0130510P.	PR 20-JUL-1999;	99US-0144632P.
PR 23-APR-1999;	99US-0130891P.	PR 21-JUL-1999;	99US-0144884P.
PR 28-APR-1999;	99US-0131449P.	PR 21-JUL-1999;	99US-0144814P.
PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145086P.
PR 30-APR-1999;	99US-0132407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-0132484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-0132485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-0132486P.	PR 22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999;	99US-0132487P.	PR 22-JUL-1999;	99US-0145192P.
PR 07-MAY-1999;	99US-0132863P.	PR 23-JUL-1999;	99US-0145145P.
PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134219P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135629P.	PR 03-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0148368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140659P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140911P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141827P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0155645P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144086P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.

PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 77.0%; Score 15.4; DB 3; Length 2132;
Best Local Similarity 94.1%; Pred. No. 9, 1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCATATCATGATGAT 19
|||||
Db 39 TCATATCATGATGAT 23

RESULF 31
AAC40444/C
ID AAC40444 standard; DNA; 2135 BP.

XX AAC40444;
XX
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 28322.

XX Hybridization assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123588P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0133048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134255P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137502P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139117P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0138763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.

PR	30-JUL-1999	99US-0144884B
PR	21-JUL-1999	99US-0144806P
PR	21-JUL-1999	99US-0145086P
PR	21-JUL-1999	99US-0145088P
PR	22-JUL-1999	99US-0145085P
PR	22-JUL-1999	99US-0145087P
PR	22-JUL-1999	99US-0145089P
PR	22-JUL-1999	99US-0145192P
PR	23-JUL-1999	99US-0145145P
PR	23-JUL-1999	99US-0145128P
PR	23-JUL-1999	99US-0145224P
PR	26-JUL-1999	99US-0145276P
PR	27-JUL-1999	99US-0145913P
PR	27-JUL-1999	99US-0145918P
PR	27-JUL-1999	99US-0145919P
PR	28-JUL-1999	99US-0145951P
PR	02-AUG-1999	99US-0146386P
PR	02-AUG-1999	99US-0146388P
PR	02-AUG-1999	99US-0146389P
PR	06-AUG-1999	99US-0147303P
PR	06-AUG-1999	99US-0147416P
PR	09-AUG-1999	99US-0147493P
PR	09-AUG-1999	99US-0147935P
PR	10-AUG-1999	99US-0148171P
PR	11-AUG-1999	99US-0148319P
PR	12-AUG-1999	99US-0148341P
PR	13-AUG-1999	99US-0148565P
PR	13-AUG-1999	99US-0148684P
PR	16-AUG-1999	99US-0149368P
PR	17-AUG-1999	99US-01494175P
PR	18-AUG-1999	99US-0149426P
PR	20-AUG-1999	99US-0149722P
PR	20-AUG-1999	99US-0149723P
PR	20-AUG-1999	99US-0149929P
PR	21-AUG-1999	99US-0149930P
PR	23-AUG-1999	99US-0149930P
PR	23-AUG-1999	99US-0149930P
PR	30-AUG-1999	99US-0151303P
PR	31-AUG-1999	99US-0151448P
PR	01-SEP-1999	99US-0151930P
PR	07-SEP-1999	99US-0152633P
PR	10-SEP-1999	99US-0153070P
PR	13-SEP-1999	99US-0153758P
PR	15-SEP-1999	99US-0154018P
PR	16-SEP-1999	99US-0154039P
PR	20-SEP-1999	99US-0154717P
PR	20-SEP-1999	99US-0155139P
PR	22-SEP-1999	99US-0155466P
PR	23-SEP-1999	99US-0155465P
PR	24-SEP-1999	99US-0156488P
PR	28-SEP-1999	99US-0156488P
PR	29-SEP-1999	99US-0156516P
PR	04-OCT-1999	99US-0157117P
PR	05-OCT-1999	99US-0157753P
PR	06-OCT-1999	99US-0157865P
PR	07-OCT-1999	99US-0158029P
PR	08-OCT-1999	99US-0158232P
PR	12-OCT-1999	99US-0158369P
PR	13-OCT-1999	99US-0159293P
PR	13-OCT-1999	99US-0159295P
PR	14-OCT-1999	99US-0159329P
PR	14-OCT-1999	99US-0159330P
PR	14-OCT-1999	99US-0159331P
PR	14-OCT-1999	99US-0159637P

PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match 77.0%; Score 15.4; DB 3; Length 2135;
Best Local Similarity 94.1%; Pred. No. 9.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy	3	TCCAACTCATGTTGAT 19
Db	39	TCCAACTCATGTTGAT 23

```

RESULT 32
AAH54374
ID AAH54374 standard; DNA; 2946 BP.
XX
AC AAH54374;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3738.
XX
KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
KM endocarditis; ds.
OS
XX Staphylococcus epidermidis.
XX OS
XX W0200134809-A2.
PN
PD 17-MAY-2001.
PF 09-NOV-2000; 2000MO-US030782.
PR 09-NOV-1999; 99US-0164258P.
PA (GLAX ) GLAXO GROUP LTD.
XX
XX Kimerly MJ;
XX WPI; 2001-316495/33.
DR
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
XX
PS Claim 8; Page 1360-1361; 218bp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AA681454 to AA683120, from Staphylococcus epidermidis. (1)
CC and (11) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce host cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC
```

CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed *S. epidermidis* genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to AAH5098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX
SQ Sequence 2946 BP; 1074 A; 481 C; 438 G; 953 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 4; Length 2946;
Best Local Similarity 94.1%; Pred. No. 9.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCGATCTCATGTTGA 18
Db 1300 GTCGATCTCATGTTGA 1316

RESULT 33
AAH54380/c
ID AAH54380 standard; DNA; 3021 BP.
XX
AC AAH54380;
XX
DT 03-SEP-2001 (first entry)
XX
DE *S. epidermidis* genomic polynucleotide sequence SEQ ID NO:3744.
XX
KW *Staphylococcus epidermidis* SRI strain; infection; diagnosis; vaccination;
XX endocarditis; ds.
XX
OS *Staphylococcus epidermidis*.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Kimerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
XX useful for vaccinating against infections, e.g. endocarditis.
XX
PS Claim 8; Page 1367-1368; 2188pp; English.

XX
CC AAH5304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AG81454 to AG83120, from *Staphylococcus epidermidis*. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the *S.*
CC *epidermidis* polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH5371 to
CC AAH5090 represent specifically claimed *S. epidermidis* genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to AAH5098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention

CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX
SQ Sequence 3021 BP; 948 A; 504 C; 458 G; 1111 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 4; Length 3021;
Best Local Similarity 94.1%; Pred. No. 9.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCGATCTCATGTTGA 18
Db 1170 GTCGATCTCATGTTGA 1154

RESULT 34
AAH54393
ID AAH54393 standard; DNA; 3391 BP.
XX
AC AAH54393;
XX
DT 03-SEP-2001 (first entry)
XX
DE *S. epidermidis* genomic polynucleotide sequence SEQ ID NO:3757.
XX
KW *Staphylococcus epidermidis* SRI strain; infection; diagnosis; vaccination;
XX endocarditis; ds.
XX
OS *Staphylococcus epidermidis*.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Kimerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
XX useful for vaccinating against infections, e.g. endocarditis.
XX
PS Claim 8; Page 1362-1383; 2188pp; English.

XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AG81454 to AG83120, from *Staphylococcus epidermidis*. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the *S.*
CC *epidermidis* polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH5371 to
CC AAH5090 represent specifically claimed *S. epidermidis* genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to AAH5098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX
SQ Sequence 3391 BP; 1291 A; 512 C; 581 G; 1007 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 3391;
Best Local Similarity 94.1%; Pred. No. 9.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTTCATCTCATGTGTA 18
|||||
Db 2802 GTTCATCTCATGTGTA 2818

RESULT 35
AAS85546
ID AAS85546 standard; cDNA; 7722 BP.

AC AAS85546;
XX
XX
DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #21350.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PsDB; ABG21359.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 1; SEQ ID NO 21350; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polynucleotide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. Ab664197-AB894564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 7722 BP; 2434 A; 1461 C; 1556 G; 2271 T; 0 U; 0 Other;

XX Query Match 77.0%; Score 15.4; DB 5; Length 7722;

Best Local Similarity 94.1%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTCATCTCATGTG 17
|||||
Db 1812 GGTCCTCTCATGTG 1828

RESULT 36
ADD46613/C
ID ADD46613 standard; DNA; 28564 BP.

XX ADD46613;

XX 29-JAN-2004 (first entry)

XX Human gene AF297093, SEQ ID NO 12296.

XX Human; ds; gene; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002MO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GENO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AF297093.

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was

CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 28564 BP; 7531 A; 5791 C; 5944 G; 9298 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 9; Length 28564;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCTACTCATGTTGA 18
DB 19285 GTCCTACTCATGTTGA 19269

RESULT 37
ADD48581/c
ID ADD48581 standard; DNA; 28564 BP.
AC ADD48581;
XX
XX 29-JAN-2004 (first entry)
DT
XX Human gene AF297093, SEQ ID NO 14287.
DE
XX Human; ds; gene; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SN; Chung.
XX
XX Homo sapiens.
OS
XX MO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX 14-AUG-2002; 2002WO-US025765.
PF
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-033347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX WPI; 2003-268312/26.
DR GENBANK; AF297093.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 28564 BP; 7531 A; 5791 C; 5944 G; 9298 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 9; Length 28564;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCTACTCATGTTGA 18
DB 19285 GTCCTACTCATGTTGA 19269

RESULT 38
AA161371/c
ID AA161371 standard; DNA; 335913 BP.
XX
XX AA161371;
AC
XX 16-OCT-2001 (first entry)
DT
XX Soybean 240017 region G3, SEQ ID NO: 2.
DE
XX
XX Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;
XX SCN resistance; rhg4; Rhg4; SCN resistant allele; plant breeding;
XX 240017 region G3; 318013 region A3; 515002 region G2; ds.
XX
XX Glycine max.
OS
XX MO200151627-A2.
PN
XX 19-JUL-2001.
PD
XX 05-JAN-2001; 2001WO-US000552.
PF
XX 07-JAN-2000; 2000US-0174880P.
PR
XX (MONS) MONSANTO CO.
XX
XX Hauge BM, Wang ML, Parsons JD, Parnell LD;
PI
XX WPI; 2001-425872/45.
DR P-PSDB; AAM42214.
XX
XX New purified nucleic acid for producing a soybean plant having soybean
PT cyst nematode resistance and for use in plant breeding programs.
PT
XX Claim 2; Page 204-400; 1353pp; English.

CC The invention relates to nucleic acid molecules from regions of the
CC soybean genome which are associated with soybean cyst nematode (SCN)
CC resistance. The nucleic acids are used to transform plants, and can
CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
CC of soybean plants and for introgressing SCN resistance or partial SCN
CC resistance into soybean plants. They can also be used in plant breeding
CC programmes. The invention also relates to proteins encoded by such
CC nucleic acid molecules, as well as antibodies capable of recognising
CC these proteins. The present sequence is a nucleic acid molecule provided
CC in the specification

XX
SQ Sequence 335913 BP; 114579 A; 53403 C; 53026 G; 114905 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 5; Length 335913;
Best Local Similarity 94.1%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGAT 19
 |||||
 DB 24565 TCCAAACTCATGTTGAT 24549

RESULT 39
 AAI61372/C
 ID AAI61372 standard; DNA; 335913 BP.
 XX
 AC AAI61372;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Soybean 240017 region G3, SEQ ID NO: 3.
 XX
 KM Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;
 KM SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;
 KM 240017 region G3; 318013 region A3; 515002 region G2; ds.
 XX
 OS Glycine max.
 XX
 PN WO200151627-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 05-JAN-2001; 2001WO-US000552.
 XX
 PR 07-JAN-2000; 2000US-0174880P.
 XX
 PA (MONS) MONSANTO CO.
 XX
 PI Hauge BM, Wang ML, Parsons JD, Parnell LD;
 DR WPI: 2001-425872/45.
 DR P-FEDB; AAM42213.
 XX
 PT New purified nucleic acid for producing a soybean plant having soybean
 PT cyst nematode resistance and for use in plant breeding programs.
 XX
 PS Claim 2; Page 400-595; 1353pp; English.

CC The invention relates to nucleic acid molecules from regions of the
 CC soybean genome which are associated with soybean cyst nematode (SCN)
 CC resistance. The nucleic acids are used to transform plants, and can
 CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
 CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
 CC of soybean plants and for introgressing SCN resistance or partial SCN
 CC resistance into soybean plants. They can also be used in plant breeding
 CC programmes. The invention also relates to proteins encoded by such
 CC nucleic acid molecules, as well as antibodies capable of recognising
 CC these proteins. The present sequence is a nucleic acid molecule provided
 CC in the specification

XX
 SQ Sequence 335913 BP; 114582 A; 53398 C; 53027 G; 114906 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 5; Length 335913;
 Best Local Similarity 94.1%; Pred. No. 1.7e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGAT 19
 |||||
 DB 24565 TCCAAACTCATGTTGAT 24549

RESULT 40
 AAC31848/C
 ID AAC31848 standard; cDNA; 308 BP.
 XX
 AC AAC31848;
 XX
 DT 06-OCT-2000 (first entry)
 XX

DE Human secreted protein 5' EST, SEQ ID NO: 35923.
 XX
 KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 1; SEQ ID NO 35923; 71pp + Sequence Listing; English.

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
 CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
 CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors

XX
 SQ Sequence 308 BP; 105 A; 54 C; 53 G; 93 T; 0 U; 3 Other;

Query Match 76.0%; Score 15.2; DB 3; Length 308;
 Best Local Similarity 85.0%; Pred. No. 8.8e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCATACATCATGTTGATG 20
 |||||
 DB 98 GATCCACATCATGTTGATG 79

RESULT 41
 AAV75448
 ID AAV75448 standard; DNA; 330 BP.
 XX
 AC AAV75448;
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE Staphylococcus aureus contig SEQ ID #1137.
 XX
 KM Computer readable medium; vaccine; S. aureus infection; immunodetection;
 KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KM skin infection; surgical wound infection; scalded skin syndrome;
 KM toxic shock syndrome; ds.
 XX
 OS Staphylococcus aureus.
 XX
 PN EP786519-A2.
 XX
 PD 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.
 PF 05-JAN-1996; 96US-0009861P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Kunsh CA, Choi GH, Barash SC, Dillon RJ, Fannon MR, Rosen CA;
 PI WPI; 1997-374922/35.
 DR
 XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
 PT stored on computer readable medium and used in the production of anti-
 PT *S.aureus* vaccines.
 PS
 PS Claim 1; Page 1850; 3271pp; English.
 CC This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the *S.aureus* DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against *S.aureus* infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC *S.aureus* in a sample. *S.aureus* is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the *S.aureus* DNA sequences contained on the computer
 CC readable medium
 SQ Sequence 330 BP; 108 A; 48 C; 54 G; 116 T; 0 U; 4 Other;
 QY Query Match 76.0%; Score 15.2; DB 2; Length 330;
 Best Local Similarity 85.0%; Pred. No. 8.9e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 DB 1 GGTCACACTCATGTTGATG 20
 284 GGACCACTCATGTTGATG 303
 RESULT 42
 AAD50058
 ID AAD50058 standard; DNA; 443 BP.
 AC AAD50058;
 XX
 XX 24-MAR-2003 (first entry)
 DT
 XX Human fatty acid elongation enzyme-like protein DNA #15.
 DE
 XX Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;
 KW long chain polyunsaturated fatty acid elongation enzyme-like protein;
 KW neurotropic; neuroprotective; antiinflammatory; de.
 XX
 OS Homo sapiens.
 XX
 XX WO200264761-A2.
 PN
 XX 22-AUG-2002.
 PD
 XX 07-FEB-2002; 2002WO-EP001260.
 PS
 XX 09-FEB-2001; 2001US-0267415P.
 PR 16-NOV-2001; 2001US-0331449P.
 PR 04-DEC-2001; 2001US-0334948P.
 XX

PA (FARB) BAYER AG.
 XX
 XX Zhu Z;
 XX
 XX WPI; 2003-040508/03.
 DR
 XX Human long chain fatty acid elongation enzyme-like protein is regulated
 PT in the treatment of diseases e.g. cancer, diabetes and CNS disorders.
 PT
 XX Disclosure; Fig 17; 175pp; English.
 PS
 XX The invention relates to an isolated polynucleotide encoding a long chain
 CC polyunsaturated fatty acid elongation enzyme-like protein polypeptide.
 CC The invention is useful in the preparation of a medicament for modulating
 CC the activity/function of long chain polyunsaturated fatty acid elongation
 CC enzyme-like protein in a disease specifically cancer, diabetes, a CNS
 CC disorder, metabolic disease, asthma or COPD. The present sequence is
 CC human long chain polyunsaturated fatty acid elongation enzyme-like
 CC protein DNA
 CC
 SQ Sequence 443 BP; 170 A; 68 C; 66 G; 139 T; 0 U; 0 Other;
 QY Query Match 76.0%; Score 15.2; DB 7; Length 443;
 Best Local Similarity 85.0%; Pred. No. 9.3e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 DB 1 GGTCACACTCATGTTGATG 20
 422 GATCCACACTCATGTTGATG 441
 RESULT 43
 AAF11349/c
 ID AAF11349 standard; cDNA; 468 BP.
 AC AAF11349;
 XX
 XX 13-MAR-2001 (first entry)
 DT
 XX *Aspergillus niger* EST SEQ ID NO:3872.
 DE
 XX Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;
 KW *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombinant;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 XX
 OS *Aspergillus niger*.
 XX
 XX WO200056762-A2.
 PN
 XX 28-SEP-2000.
 PD
 XX 22-MAR-2000; 2000WO-US007781.
 PF
 XX 22-MAR-1999; 99US-00273623.
 PR
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 PA
 XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 PI WPI; 2000-594572/56.
 DR
 XX Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags.
 XX
 XX Claim 87; Page 1730; 3161pp; English.
 PS
 XX The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (PF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal

CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the PF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring the
CC global expression of genes from FF cells allows the production potential
CC of the microorganisms to be improved. New genes may be discovered,
CC possible functions of unknown open reading frames can be identified and
CC gene copy number variation and stability can be monitored. The expression
CC of genes can be used to study how PF cells adapt to changes in culture
CC conditions, environmental stress, spore morphogenesis, recombination,
CC metabolic or catabolic pathway engineering. Using ESTs provides several
CC advantages over genomic or random cDNA clones including elimination of
CC redundancy as one spot on an array equals one gene or open reading frame,
CC and organisation of the microarray based on function of the gene
CC products to facilitate analysis of the results. AAF07478 to AAF11247
CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents
CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from
CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from
CC *Trichoderma reesei*, which are all specifically claimed in the present
CC invention

XX SQ Sequence 468 BP; 109 A; 129 C; 118 G; 111 T; 0 U; 1 Other;

Query Match 76.0%; Score 15.2; DB 3; Length 468;

Best Local Similarity 85.0%; Pred. No. 9.3e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCATCATCATGTTGATG 20

102 GGTTCATACCATGTTGCTG 83

RESULT 44

AD50056

AD50056 standard; DNA; 472 BP.

AC AAD50056;

DT 24-MAR-2003 (first entry)

DE Human fatty acid elongation enzyme-like protein DNA #13.

XX Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;

KW long chain polyunsaturated fatty acid elongation enzyme-like protein;

KW neurotropic; neuroprotective; antiinflammatory; de.

XX Homo sapiens.

PN WO200264761-A2.

PD 22-AUG-2002.

PF 07-FEB-2002; 2002WO-EP001260.

PR 09-FEB-2001; 2001US-0267415P.

PR 16-NOV-2001; 2001US-0331449P.

PR 04-DEC-2001; 2001US-0334948P.

PA (FARB) BAYER AG.

PI Zhu Z;

DR MPI; 2003-040508/03.

XX Human long chain fatty acid elongation enzyme-like protein is regulated
PT in the treatment of diseases e.g. cancer, diabetes and CNS disorders.

XX Disclosure; Fig 15, 175pp; English.

CC The invention relates to an isolated polynucleotide encoding a long chain
CC polyunsaturated fatty acid elongation enzyme-like protein polypeptide.

CC The invention is useful in the preparation of a medicament for modulating

CC the activity/function of long chain polyunsaturated fatty acid elongation
CC enzyme-like protein in a disease specifically cancer, diabetes, a CNS
CC disorder, metabolic disease, asthma or COPD. The present sequence is
CC human long chain polyunsaturated fatty acid elongation enzyme-like
CC protein DNA

XX SQ Sequence 472 BP; 174 A; 72 C; 74 G; 152 T; 0 U; 0 Other;

Query Match 76.0%; Score 15.2; DB 7; Length 472;

Best Local Similarity 85.0%; Pred. No. 9.4e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCATCATCATGTTGATG 20

419 GATCCACATCATGTTGATG 438

RESULT 45

ABK69101/C

ID ABK69101 standard; cDNA; 479 BP.

AC ABK69101;

DT 02-JUL-2002 (first entry)

DE DNA encoding human secreted protein, SEQ ID No 25.

KW Human; secreted protein; autoimmune disease; rheumatoid arthritis;

KW hyperproliferative disorder; neoplasm; breast; liver; ischaemia;

KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;

KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;

KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;

KW epithelial cell proliferation; organ transplantation; food additive;

KW food storage; gene; ss.

XX Homo sapiens.

PN WO200224721-A1.

PD 28-MAR-2002.

PF 09-JAN-2001; 2001WO-US000544.

PR 20-SEP-2000; 2000US-0234211P.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Komatsoulis GA, Baker KP, Rosen CA, Birse CE, Soppet DR;

PI Olsen HS, Ni J, Fiscella M, Moore PA, Wei P, Ebner R, Duan DR;

PI Shi Y, Choi GH;

DR MPI; 2002-330012/36.

DR P-PSDB; AAU96180.

PT Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.

XX Claim 1; Page 462; 562pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) encoding a
CC human secreted protein (II). (I) and (II) are used to prevent, treat or
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,

CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in

CC diagnosing a pathological condition or susceptibility to a pathological

CC condition. The antibodies to (II) can also be used in alleviating

CC symptoms associated with the disorders and in diagnostic immunoassays

CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).

CC Disorders which are diagnosed or treated include autoimmune diseases e.g.

CC rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the

CC breast or liver, cardiovascular disorders e.g. cardiac arrest,

CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous

CC system disorders e.g. Alzheimer's disease, infections caused by bacteria,

CC viruses and fungi and ocular disorders e.g. corneal infection. The

CC polypeptides can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues and in chemotaxis. The polypeptides can also be
CC used as a food additive or preservative to increase or decrease storage
CC capabilities. ABK69078-ABK69143 represent human secreted protein coding
CC sequences, PCR primers and related sequences used in cloning and
CC expression of the secreted proteins described in examples of the
CC invention
XX
SQ Sequence 479 BP, 106 A, 169 C, 107 G, 97 T, 0 U, 0 Other;

Query Match 76.0%; Score 15.2; DB 6; Length 479;
Best Local Similarity 85.0%; Pred. No. 9.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCACATCATGTTGATG 20
DB 373 GGTCACATCATGTTGATG 354

RESULT 46
AAD50057/c
ID AAD50057 standard; DNA; 499 BP.
XX
AC AAD50057;
XX
DT 24-MAR-2003 (first entry)
XX
DE Human fatty acid elongation enzyme-like protein DNA #14.
XX
KW Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;
KW long chain polyunsaturated fatty acid elongation enzyme-like protein;
KW neurotropic; neuroprotective; antiinflammatory; de.
XX
OS Homo sapiens.
XX
PN WO200264761-A2.
XX
PD 22-AUG-2002.
XX
PF 07-FEB-2002; 2002WO-EP001260.
XX
PR 09-FEB-2001; 2001US-0267415P.
PR 16-NOV-2001; 2001US-0331449P.
PR 04-DEC-2001; 2001US-0334948P.
XX
PA (FARB) BAYER AG.
XX
PI Zhu Z;
XX
DR WPI; 2003-040508/03.
XX
PT Human long chain fatty acid elongation enzyme-like protein is regulated
PT in the treatment of diseases e.g. cancer, diabetes and CNS disorders.
XX
PS Disclosure; Fig 16; 175pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a long chain
CC polyunsaturated fatty acid elongation enzyme-like protein polypeptide.
CC The invention is useful in the preparation of a medicament for modulating
CC the activity/function of long chain polyunsaturated fatty acid elongation
CC enzyme-like protein in a disease specifically cancer, diabetes, a CNS
CC disorder, metabolic disease, asthma or COPD. The present sequence is
CC human long chain polyunsaturated fatty acid elongation enzyme-like
CC protein DNA
XX
SQ Sequence 499 BP, 174 A, 97 C, 82 G, 145 T, 0 U, 1 Other;

Query Match 76.0%; Score 15.2; DB 7; Length 499;
Best Local Similarity 85.0%; Pred. No. 9.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCACATCATGTTGATG 20
DB 304 GATCCACATCATGTTGATG 285

RESULT 47
ABK69131/c
ID ABK69131 standard; cDNA; 547 BP.
XX
AC ABK69131;
XX
DT 02-JUL-2002 (first entry)
XX
DE DNA encoding human secreted protein, SEQ ID No 55.
XX
KW Human; secreted protein; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; breast; liver; ischaemia;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;
KW epithelial cell proliferation; organ transplantation; food additive;
KW food storage; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200224721-A1.
XX
PD 28-MAR-2002.
XX
PF 09-JAN-2001; 2001WO-US000544.
XX
PR 20-SEP-2000; 2000US-0234211P.
XX
PA (HDNA-) HUMAN GENOME SCI INC.
XX
PI Komatsu Is GA, Baker KP, Rosen CA, Birse CE, Soppet DR;
PI Olsen HS, Ni J, Fiscella M, Moore PA, Wei P, Ebner R, Duan DR;
PI Shi Y, Choi GH;
XX
DR WPI; 2002-330012/36.
DR P-PSDB; AAU96210.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.
XX
PS Claim 1; Page 484; 562pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) encoding a
CC human secreted protein (II). (I) and (II) are used to prevent, treat or
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to (II) can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases e.g.
CC rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the
CC breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous
CC system disorders e.g. Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi and ocular disorders e.g. corneal infection. The
CC polypeptides can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues and in chemotaxis. The polypeptides can also be
CC used as a food additive or preservative to increase or decrease storage
CC capabilities. ABK69078-ABK69143 represent human secreted protein coding
CC sequences, PCR primers and related sequences used in cloning and
CC expression of the secreted proteins described in examples of the
CC invention
XX
SQ Sequence 547 BP, 119 A, 185 C, 121 G, 117 T, 0 U, 5 Other;

Query Match 76.0%; Score 15.2; DB 6; Length 547;
 Best Local Similarity 85.0%; Pred. No. 9.5e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTTGATG 20
 DB 366 GGTCCATCTCATGTTGATG 347

RESULT 48
 AAD50054
 ID AAD50054 standard; DNA; 548 BP.
 XX
 AC AAD50054;
 XX
 DT 24-MAR-2003 (first entry)
 DE Human fatty acid elongation enzyme-like protein DNA #11.
 XX
 KW Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;
 KM long chain polyunsaturated fatty acid elongation enzyme-like protein;
 KM nootropic; neuroprotective; antiinflammatory; ds.
 XX
 OS Homo sapiens.
 XX
 PI WO200264761-A2.
 XX
 PD 22-AUG-2002.
 XX
 PS 07-FEB-2002; 2002WO-EP001260.
 XX
 PR 09-FEB-2001; 2001US-0267415P.
 XX
 PR 16-NOV-2001; 2001US-0331449P.
 PR 04-DEC-2001; 2001US-0334948P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Zhu Z;
 XX
 DR WPI; 2003-040508/03.
 XX
 PT Human long chain fatty acid elongation enzyme-like protein is regulated
 PT in the treatment of diseases e.g. cancer, diabetes and CNS disorders.
 XX
 PS Disclosure; Fig 13; 175pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a long chain
 CC polyunsaturated fatty acid elongation enzyme-like protein polypeptide.
 CC The invention is useful in the preparation of a medicament for modulating
 CC the activity/function of long chain polyunsaturated fatty acid elongation
 CC enzyme-like protein in a disease specifically cancer, diabetes, a CNS
 CC disorder, metabolic disease, asthma or COPD. The present sequence is
 CC human long chain polyunsaturated fatty acid elongation enzyme-like
 CC protein DNA
 XX
 SQ Sequence 548 BP; 192 A; 84 C; 86 G; 186 T; 0 U; 0 Other;
 XX
 Query Match 76.0%; Score 15.2; DB 7; Length 548;
 Best Local Similarity 85.0%; Pred. No. 9.5e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTTGATG 20
 DB 422 GATCCACATCATGTTGATG 441

RESULT 49
 AAD50053
 ID AAD50053 standard; DNA; 578 BP.
 XX
 AC AAD50053;
 XX
 DT 24-MAR-2003 (first entry)

XX
 DE Human fatty acid elongation enzyme-like protein DNA #10.
 XX
 KW Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;
 KM long chain polyunsaturated fatty acid elongation enzyme-like protein;
 KM nootropic; neuroprotective; antiinflammatory; ds.
 XX
 OS Homo sapiens.
 XX
 PI WO200264761-A2.
 XX
 PD 22-AUG-2002.
 XX
 PS 07-FEB-2002; 2002WO-EP001260.
 XX
 PR 09-FEB-2001; 2001US-0267415P.
 XX
 PR 16-NOV-2001; 2001US-0331449P.
 PR 04-DEC-2001; 2001US-0334948P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Zhu Z;
 XX
 DR WPI; 2003-040508/03.
 XX
 PT Human long chain fatty acid elongation enzyme-like protein is regulated
 PT in the treatment of diseases e.g. cancer, diabetes and CNS disorders.
 XX
 PS Disclosure; Fig 12; 175pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a long chain
 CC polyunsaturated fatty acid elongation enzyme-like protein polypeptide.
 CC The invention is useful in the preparation of a medicament for modulating
 CC the activity/function of long chain polyunsaturated fatty acid elongation
 CC enzyme-like protein in a disease specifically cancer, diabetes, a CNS
 CC disorder, metabolic disease, asthma or COPD. The present sequence is
 CC human long chain polyunsaturated fatty acid elongation enzyme-like
 CC protein DNA
 XX
 SQ Sequence 578 BP; 199 A; 87 C; 87 G; 205 T; 0 U; 0 Other;
 XX
 Query Match 76.0%; Score 15.2; DB 7; Length 578;
 Best Local Similarity 85.0%; Pred. No. 9.6e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTTGATG 20
 DB 439 GATCCACATCATGTTGATG 458

RESULT 50
 AAF14157
 ID AAF14157 standard; cDNA; 685 BP.
 XX
 AC AAF14157;
 XX
 DT 13-MAR-2001 (first entry)
 DE Aspergillus oryzae EST SEQ ID NO:6680.
 XX
 KW Multiple gene expression; filamentous fungal cell; EST;
 KM expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KM Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KM culture condition; environmental stress; spore morphogenesis;
 KM metabolic pathway engineering; catabolic pathway engineering; ss.
 XX
 OS Aspergillus oryzae.
 XX
 PN WO200056762-A2.
 XX
 PD 28-SEP-2000.
 XX
 PF 22-MAR-2000; 2000WO-US007781.

XX 22-MAR-1999; 99US-00273623 .
PR

XX (NOVO) NOVO NORDISK BIOTECH INC
PA (NOVO) NOVO NORDISK AS.
PA

PA (NOVO) NOVO NORDISK A.S.

PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

DR WPI; 2000-594572/56.

PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags.

PT substrate of expressed sequence tags.

PT substrate of expressed sequence tags.

PS Claim 88; Page 2728-2729; 3161pp; English

The present invention describes a method for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, possible functions of unknown open reading frames can be identified and gene copy number variation and stability can be monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organisation of the microarrays based on function of the gene. CC products to facilitate analysis of the results. AAFF07478 to AAFF12447 CC represents ESTs from *Fusarium venenatum*; AAFF1248 to AAFF1853 represents ESTs from *Aspergillus niger*; AAFF1854 to AAFF14878 represents ESTs from *Aspergillus oryzae*; and AAFF14879 to AAFF15337 represents ESTs from *Trichoderma reesei*, which are all specifically claimed in the present invention.

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

CC invention

Sequence 685 BP; 169 A; 178 C; 174 G; 164 T; 0 U; 0 Other;

Query Match	76.0%;	Score 15.2;	DB 3;	Length 685;
-------------	--------	-------------	-------	-------------

Best Local Similarity 85.0%; Pred. No. 9.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY      1  GGTCCATACCTCATGTTGATG  20
          |||||
Db      353 GGTCCATACCTCTTGATGCTG  372

```

Db 353 GGTCCTACTCTTGATGCTG 372

Search completed: August 17, 2004, 14:36:28
Job time : 402 secs

Job time : 402 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 14:16:42 ; Search time 2541 Seconds
(without alignments)
235.043 Million cell updates/sec

Title: US-09-825-489-3

Perfect score: 20

Sequence: 1 ggtccatcatcatgtcatg 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues 55026578

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 800 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estmu:*
4: em_estov:*
5: em_estro:*
6: em_estro:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
20	100.0	219	12	BG201597	RST20938
20	100.0	221	12	BG212172	RST31643
20	100.0	352	12	BG185555	RST4504 A
20	100.0	359	14	CB852193	UI-CF-FNO

5	100.0	370	12	BM796479	BM796479 K-EST0079
6	100.0	408	9	AV687532	AV687532 AV687532
7	100.0	408	9	AV690967	AV690967 AV690967
8	100.0	437	9	AI961077	AI961077 wq44g01.x
9	100.0	461	13	BM679004	BM679004 UI-CF-D01
10	100.0	486	9	AA628925	AA628925 od81a10.8
11	100.0	487	12	BI793276	BI793276 1e50b09.y
12	100.0	506	9	AA167098	AA167098 2p05g10.x
13	100.0	510	13	B0186624	B0186624 UI-E-EJ1-
14	100.0	513	9	AA165911	AA165911 2p05g10.8
15	100.0	515	13	B0184648	B0184648 UI-E-EJ1-
16	100.0	534	10	AM890854	AM890854 RCS-NT005
17	100.0	587	14	CB048344	CB048344 NISC_9104
18	100.0	591	14	CB048343	CB048343 NISC_9104
19	100.0	592	10	AM979099	AM979099 EST391209
20	100.0	595	12	BM739320	BM739320 K-EST0008
21	100.0	598	12	BI792983	BI792983 1e50b09.x
22	100.0	627	10	BF446397	BF446397 7p36g05.x
23	100.0	638	12	BM853017	BM853017 K-EST0134
24	100.0	639	12	BI911749	BI911749 603065343
25	100.0	672	14	CB305774	CB305774 UI-CF-PN1
26	100.0	678	12	BM920900	BM920900 UI-H-DL1-
27	100.0	705	14	CA450156	CA450156 UI-CF-FNO
28	100.0	748	13	BQ772342	BQ772342 UI-H-EJ1-
29	100.0	790	12	BG186679	BG186679 RST5654 A
30	100.0	809	29	AY414079	AY414079 Pan trogl
31	100.0	813	29	AY414078	AY414078 Homo sapi
32	100.0	873	9	AL545429	AL545429 AL545429
33	100.0	1059	12	BM468458	BM468458 AGNCNCURT
34	100.0	1114	9	AL573331	AL573331 AL573331
35	100.0	1129	9	AL548961	AL548961 AL548961
36	100.0	1174	13	B0879392	B0879392 AGNCNCURT
37	98.0	1201	9	AL531628	AL531628 AL531628
38	95.0	640	9	AI990924	AI990924 ws24f07.x
39	92.0	467	10	BE763751	BE763751 RCS-NT005
40	92.0	568	13	BQ321319	BQ321319 QVA-CF049
41	92.0	625	12	BJ057361	BJ057361 BJO57361
42	92.0	711	9	AT634902	AT634902 t274c02.x
43	92.0	854	12	BG187563	BG187563 RST6691 A
44	92.0	857	9	AL571156	AL571156 AL571156
45	92.0	1024	28	CC448351	CC448351 ZMMBNC032
46	87.0	420	10	AM646323	AM646323 cm63e10.w
47	87.0	492	12	BJ083380	BJ083380 BJO83380
48	87.0	587	9	AL637062	AL637062 AL637062
49	87.0	828	14	CD434346	CD434346 ELO1N0323
50	85.0	808	12	BI261004	BI261004 602972177
51	85.0	198	29	CG932703	CG932703 MBHEW59TF
52	84.0	198	29	CG486388	CG486388 OSTR21372
53	84.0	294	12	BI297979	BI297979 UI-R-CV2-
54	84.0	348	12	BI297518	BI297518 UI-R-CV2-
55	84.0	360	12	BM293462	BM293462 C0304G12-
56	84.0	365	9	AA927997	AA927997 on79h12.8
57	84.0	390	12	BQ031452	BQ031452 UI-1-CF0-
58	84.0	436	10	BF417305	BF417305 UI-R-CNO-
59	84.0	500	12	BI300394	BI300394 UI-R-CV2-
60	84.0	518	28	AZ928003	AZ928003 UI-R-CV2-
61	84.0	520	29	CG638173	CG638173 OST367192
62	84.0	531	29	CG112083	CG112083 PUIFX20TD
63	84.0	563	14	CA351074	CA351074 622055 NC
64	84.0	597	13	BM243041	BM243041 BM243041
65	84.0	598	13	B0207028	B0207028 UI-R-D21-
66	84.0	616	14	CA381609	CA381609 661145 NC
67	84.0	632	14	CD292702	CD292702 STRPUS36.
68	84.0	635	14	CA379904	CA379904 659105 NC
69	84.0	636	9	AV864708	AV864708 AV864708
70	84.0	657	29	AG115048	AG115048 Pan trogl
71	84.0	667	10	BF115636	BF115636 7b92907.x
72	84.0	669	13	BM261667	BM261667 BM261667
73	84.0	679	14	CA231933	CA231933 SCQSRFL303
74	84.0	681	13	BM270720	BM270720 BM270720
75	84.0	699	13	BX075094	BX075094 BX075094
76	84.0	700	14	CA356753	CA356753 629000 NC
77	84.0	722	13	CA060694	CA060694 bsa1rgb53

78	16.8	84.0	730	12	BI296724	UI-R-CV2-	BI296724	UI-R-CV2-	15.8	79.0	404	9	A1446339	A1446339	tj20905.x
79	16.8	84.0	757	14	CB511930	gsaltrgb54	CB511930	gsaltrgb54	15.8	79.0	411	9	A1624772	A1624772	ts44c12.x
80	16.8	84.0	789	28	BZ982973	PUGGR07TB	BZ982973	PUGGR07TB	15.8	79.0	413	9	AL544616	AL544616	AL544616
81	16.8	84.0	811	29	CG555961	OGMCH90TV	CG555961	OGMCH90TV	15.8	79.0	415	10	H05220	H05220	y185d09..81
82	16.8	84.0	819	29	AXY14080	Mus_muscu	AXY14080	Mus_muscu	15.8	79.0	419	10	AM273418	AM273418	xc39c09.x
83	16.8	84.0	830	28	BH710049	BOMOU27TR	BH710049	BOMOU27TR	15.8	79.0	443	13	BO973507	BO973507	OH13A16.
84	16.8	84.0	853	28	BH086317	RPCT-24-2	BH086317	RPCT-24-2	15.8	79.0	444	9	A1490900	A1490900	ESR724609
85	16.8	84.0	870	13	BUR40815	AGENCOURT	BUR40815	AGENCOURT	15.8	79.0	444	9	A1564466	A1564466	tg57f03.x
86	16.8	84.0	984	29	CNS00K6D	AL0777281	CNS00K6D	AL0777281	15.8	79.0	450	10	BF894540	BF894540	RC1-MT013
87	16.8	84.0	1682	10	BF572295	Drosophila1	BF572295	Drosophila1	15.8	79.0	451	28	AQ917107	AQ917107	T233136b
88	16.4	82.0	295	10	BB286508	BB286508	BB286508	BB286508	15.8	79.0	452	14	N25249	N25249	yx74d06..81
89	16.4	82.0	363	14	CAS54445	C08B5G06-	CAS54445	C08B5G06-	15.8	79.0	453	9	AA805564	AA805564	oc24d07..8
90	16.4	82.0	400	28	AQ011724	HS_-2258_A	AQ011724	HS_-2258_A	15.8	79.0	456	10	AM195941	AM195941	X185a12.x
91	16.4	82.0	400	28	AQ085335	HS_2164_B	AQ085335	HS_2164_B	15.8	79.0	457	14	CD283195	CD283195	G38941..42
92	16.4	82.0	475	28	BH14540	CH230-153	BH14540	CH230-153	15.8	79.0	461	28	BH883167	BH883167	hw42fD4..9
93	16.4	82.0	513	28	BH660918	UP_293-15	BH660918	UP_293-15	15.8	79.0	469	10	BF435867	BF435867	nab42d11..
94	16.4	82.0	522	29	BKX52083	Arabidops	BKX52083	Arabidops	15.8	79.0	470	14	H02898	H02898	yj39e03..81
95	16.4	82.0	602	29	CBS24885	L1gr-g8s-	CBS24885	L1gr-g8s-	15.8	79.0	471	9	AA743710	AA743710	ny92f04..8
96	16.4	82.0	620	14	CAS48088	C0801A09-	CAS48088	C0801A09-	15.8	79.0	475	12	BM582117	BM582117	RK20037..5
97	16.4	82.0	626	28	AQ286592	mgxb0017C	AQ286592	mgxb0017C	15.8	79.0	477	10	AM570406	AM570406	g322h03..y
98	16.4	82.0	668	28	AZ333406	IM0062C02	AZ333406	IM0062C02	15.8	79.0	477	28	AQ814813	AQ814813	HS_-5258_B
99	16.4	82.0	672	10	BB053667	BB053667	BB053667	BB053667	15.8	79.0	479	9	AA725575	AA725575	a124d11..8
100	16.4	82.0	702	28	AZ434015	1M0220009	AZ434015	1M0220009	15.8	79.0	481	10	BE074235	BE074235	QV3-BT056
101	16.4	82.0	775	14	CA429356	UI-R-FH1-	CA429356	UI-R-FH1-	15.8	79.0	483	28	AQ689450	AQ689450	tlpxb0079N
102	16.4	82.0	775	28	BH483076	BOGRD77TF	BH483076	BOGRD77TF	15.8	79.0	485	12	BM144489	BM144489	TCAPAD11
103	16.4	82.0	779	29	BX138283	Danilo_rer	BX138283	Danilo_rer	15.8	79.0	486	12	BM765921	BM765921	K-EST052
104	16.4	82.0	805	28	BH363463	CH230-5N8	BH363463	CH230-5N8	15.8	79.0	487	12	BI786529	BI786529	sa150b09..
105	16.4	82.0	824	28	AZ174590	SP_0130_A	AZ174590	SP_0130_A	15.8	79.0	507	9	A1889130	A1889130	tw36h12..x
106	16.4	82.0	831	29	CCS11958	CH240_355	CCS11958	CH240_355	15.8	79.0	508	14	CA735557	CA735557	wp1b..pk0
107	16.4	82.0	914	13	BU774772	60352350	BU774772	60352350	15.8	79.0	512	9	A1959175	A1959175	fd07g05..y
108	16.4	82.0	958	28	AQ743158	HS_-5384_B	AQ743158	HS_-5384_B	15.8	79.0	512	28	AZ688834	AZ688834	RPCT-23-2
109	16.4	82.0	1002	10	BE616227	601279242	BE616227	601279242	15.8	79.0	514	14	CB085136	CB085136	h427d12..b
110	16.4	82.0	2320	11	AK032870	Mus_muscu	AK032870	Mus_muscu	15.8	79.0	516	28	B2916626	B2916626	CH240_60J
111	16.4	82.0	3553	11	AK030322	Mus_muscu	AK030322	Mus_muscu	15.8	79.0	518	10	BE835251	BE835251	RC5-FM002
112	16	80.0	480	14	CP931227	CF--06-R-	CP931227	CF--06-R-	15.8	79.0	523	10	AM270635	AM270635	XB83-c03..x
113	16	80.0	575	28	AZ019301	RPCT-23-3	AZ019301	RPCT-23-3	15.8	79.0	526	28	BZ221678	BZ221678	XB83-c03..x
114	16	80.0	576	12	BJ053329	BJ053329	BJ053329	BJ053329	15.8	79.0	527	13	BU713949	BU713949	SJABAXA07
115	16	80.0	633	10	BF936583	EST1459638	BF936583	EST1459638	15.8	79.0	533	13	BU950227	BU950227	1o78c09.x
116	16	80.0	863	29	CG857171	ZMMBB026	CG857171	ZMMBB026	15.8	79.0	537	12	BI887105	BI887105	2P637-1-0
117	15.8	79.0	204	10	BE243298	TCAPAD09	BE243298	TCAPAD09	15.8	79.0	541	9	AV914909	AV914909	AV914909
118	15.8	79.0	207	14	H46770	yo18g05..81	H46770	yo18g05..81	15.8	79.0	546	13	BU927775	BU927775	AGENCOURT
119	15.8	79.0	264	12	BM15116	TCBAP1D61	BM15116	TCBAP1D61	15.8	79.0	548	12	BI498599	BI498599	sa116c04..
120	15.8	79.0	270	14	CB496164	omykr000	CB496164	omykr000	15.8	79.0	552	9	A1957898	A1957898	fd07g05..x
121	15.8	79.0	270	28	BZ215304	CH230-332	BZ215304	CH230-332	15.8	79.0	554	12	BG524360	BG524360	41-72_Ste
122	15.8	79.0	273	10	AM367203	MR0-HT016	AM367203	MR0-HT016	15.8	79.0	557	9	AL045488	AL045488	DKEP24341
123	15.8	79.0	275	9	AA665903	ob50g05..8	AA665903	ob50g05..8	15.8	79.0	560	28	AQ690181	AQ690181	tlpxb0081D
124	15.8	79.0	275	10	AM196133	xm05a05..x	AM196133	xm05a05..x	15.8	79.0	564	10	AM468256	AM468256	h8c0c04..x
125	15.8	79.0	298	13	BO630957	1114c10..y	BO630957	1114c10..y	15.8	79.0	564	14	CP032818	CP032818	OCB8e02..y
126	15.8	79.0	297	14	CF087871	QHM17B22..	CF087871	QHM17B22..	15.8	79.0	574	10	BF382660	BF382660	601816872
127	15.8	79.0	297	14	R24247	yh30b02..81	R24247	yh30b02..81	15.8	79.0	574	29	CE756403	CE756403	L1gr-g8s-
128	15.8	79.0	299	28	CC015488	PUDP48TD	CC015488	PUDP48TD	15.8	79.0	579	29	CE439041	CE439041	L1gr-g8s-
129	15.8	79.0	313	12	BMS10485	1j46c07..x	BMS10485	1j46c07..x	15.8	79.0	584	10	BF464803	BF464803	NFO47B08E
130	15.8	79.0	314	13	BO582223	1114c10..x	BO582223	1114c10..x	15.8	79.0	587	13	EX487241	EX487241	DKEP26860
131	15.8	79.0	317	12	BMS11316	1j46c07..y	BMS11316	1j46c07..y	15.8	79.0	589	28	AQ986586	AQ986586	RPCT-23-3
132	15.8	79.0	327	13	BU918258	5021-37_M	BU918258	5021-37_M	15.8	79.0	590	28	AZ421903	AZ421903	1M0200C24
133	15.8	79.0	329	14	R91378	BM108121.5	R91378	BM108121.5	15.8	79.0	592	14	CB163502	CB163502	K-BST0229
134	15.8	79.0	343	12	BI608116	BM108121.5	BI608116	BM108121.5	15.8	79.0	593	28	AZ406088	AZ406088	1M0175M09
135	15.8	79.0	347	14	N20854	yx51g11..81	N20854	yx51g11..81	15.8	79.0	599	10	AM673731	AM673731	ba57c01..x
136	15.8	79.0	358	14	D62710	HUM317H03B	D62710	HUM317H03B	15.8	79.0	599	28	AQ438485	AQ438485	HS_5115_B
137	15.8	79.0	362	10	AM651621	ba60b07..x	AM651621	ba60b07..x	15.8	79.0	600	28	BZ371495	BZ371495	1e37b12..b
138	15.8	79.0	362	29	CG943525	MBEGJ26TR	CG943525	MBEGJ26TR	15.8	79.0	600	28	BZ375237	BZ375237	1e37b12..g
139	15.8	79.0	365	28	AQ015910	CIT-HSP-2	AQ015910	CIT-HSP-2	15.8	79.0	605	28	AZ257832	AZ257832	RPCT-23-1
140	15.8	79.0	368	9	AL080098	DKEP2586L	AL080098	DKEP2586L	15.8	79.0	611	14	CB483863	CB483863	cc1uwbh0
141	15.8	79.0	368	10	AM512370	ku56a06..x	AM512370	ku56a06..x	15.8	79.0	615	28	B74016	B74016	T72H20TR.TA
142	15.8	79.0	383	12	BM782600	K-EST0059	BM782600	K-EST0059	15.8	79.0	620	14	CF307533	CF307533	HDA1--06-
143	15.8	79.0	384	29	CG805486	ZMMBBc047	CG805486	ZMMBBc047	15.8	79.0	622	14	CF807709	CF807709	PHB027XI
144	15.8	79.0	388	9	A1539027	ta99c01..x	A1539027	ta99c01..x	15.8	79.0	631	13	BQ410891	BQ410891	GA_Ed003
145	15.8	79.0	388	9	A1539064	tw42f11..x	A1539064	tw42f11..x	15.8	79.0	635	14	CD014962	CD014962	hac27A10..
146	15.8	79.0	388	12	BI055583	MR3-GN034	BI055583	MR3-GN034	15.8	79.0	638	28	AQ602236	AQ602236	HS_2111_B
147	15.8	79.0	389	14	H41924	yo60f08..81	H41924	yo60f08..81	15.8	79.0	645	14	CB688581	CB688581	CEST-27-D
148	15.8	79.0	391	14	N20864	yx52c11..81	N20864	yx52c11..81	15.8	79.0	653	13	BY722968	BY722968	RPCT-23-1
149	15.8	79.0	395	14	CF186400	44j23j2..f	CF186400	44j23j2..f	15.8	79.0	658	28	AZ282626	AZ282626	oem96a02..
150	15.8	79.0	398	12	BI074444	IPI_15_D0	BI074444	IPI_15_D0	15.8	79.0	661	28	BZ006698	BZ006698	

C 224	15.8	79.0	666	29	CG760052	ZMMBBD031	C 297	15.4	77.0	352	12	BI771144	BI771144	EBrc03_SQ
C 225	15.8	79.0	668	14	CD079743	MA3_-99990	298	15.4	77.0	365	28	A2255310	A2255310	RPCI-23-9
C 226	15.8	79.0	673	14	CB445568	669632_MA	C 299	15.4	77.0	372	9	AA396945	AA396945	mr1f10..x
C 227	15.8	79.0	673	14	CB449974	704590_MA	300	15.4	77.0	376	9	AI851749	AI851749	UI-M-BHD-
C 228	15.8	79.0	675	29	CG855018	ZMMBBD022	C 301	15.4	77.0	380	14	CF506831	CF506831	USDA-EP_1
C 229	15.8	79.0	676	12	BM468848	AGENCOCRT	C 302	15.4	77.0	381	9	AI216097	AI216097	gm4b003..x
C 230	15.8	79.0	679	29	CE661268	ligr-g88-	C 303	15.4	77.0	386	12	BG232692	BG232692	pa78g04..y
C 231	15.8	79.0	695	28	BH506396	BH506396	C 304	15.4	77.0	392	13	BY605976	BY605976	BY605976
C 232	15.8	79.0	699	9	AV938999	AV938999	C 305	15.4	77.0	393	9	BY766819	BY766819	ah37a12..s
C 233	15.8	79.0	699	14	CB688262	CEST-26-B	C 306	15.4	77.0	393	9	AA209714	AA209714	mw72b02..x
C 234	15.8	79.0	704	12	BI931092	EST550981	C 307	15.4	77.0	400	14	CF506703	CF506703	USDA-EP_1
C 235	15.8	79.0	707	13	BM038685	BM038685	C 308	15.4	77.0	407	9	AI655581	AI655581	lc27h12..x
C 236	15.8	79.0	711	14	CA423177	UI-H-FE1-	C 309	15.4	77.0	407	10	BF061808	BF061808	7k67C02..x
C 237	15.8	79.0	714	14	CA413785	UI-H-EZ0-	C 310	15.4	77.0	407	13	BY228504	BY228504	BY228504
C 238	15.8	79.0	714	28	BH578920	BH578920	C 311	15.4	77.0	421	10	R49177	R49177	Y958a10..bl
C 239	15.8	79.0	716	28	BH941998	odf46h06.	C 312	15.4	77.0	428	29	EX189233	EX189233	Danio_rer
C 240	15.8	79.0	722	12	BIJ51331	BIJ51331	C 313	15.4	77.0	429	9	AA387236	AA387236	vc21g11..x
C 241	15.8	79.0	722	12	BG468373	60250983	C 314	15.4	77.0	435	10	BE944649	BE944649	UI-M-BH3-
C 242	15.8	79.0	728	10	BE381698	601272121	C 315	15.4	77.0	441	9	AA013929	AA013929	mh24g08..x
C 243	15.8	79.0	736	29	CEA17738	ligr-g88-	C 316	15.4	77.0	449	13	BI113931	BI113931	mh24g08..x
C 244	15.8	79.0	738	28	BZ523914	BORAF75TF	C 317	15.4	77.0	450	14	CD372378	CD372378	UI-R-GD0-
C 245	15.8	79.0	739	12	BG615504	602642809	C 318	15.4	77.0	452	9	AA881328	AA881328	vx12h08..x
C 246	15.8	79.0	742	28	BH650112	BOMBL43TF	C 319	15.4	77.0	453	28	AQ117439	AQ117439	2191_A
C 247	15.8	79.0	747	29	CE567553	ligr-g88-	C 320	15.4	77.0	455	13	BY465273	BY465273	BY465273
C 248	15.8	79.0	753	14	CD743096	UI-H-Ftl-	C 321	15.4	77.0	456	9	AA013931	AA013931	mh24h08..x
C 249	15.8	79.0	755	14	CF738663	UI-M-HD0-	C 322	15.4	77.0	474	10	BF660500	BF660500	ma992f12.
C 250	15.8	79.0	767	14	CB229722	AGENCOCRT	C 323	15.4	77.0	476	14	CF156184	CF156184	ma992f12.
C 251	15.8	79.0	767	28	BH650336	BH650336	C 324	15.4	77.0	477	14	CF158701	CF158701	B0663A01-
C 252	15.8	79.0	769	28	BH490051	BOHGU57TR	C 325	15.4	77.0	477	14	CF158761	CF158761	B0663G02-
C 253	15.8	79.0	783	28	BH435182	BH435182	C 326	15.4	77.0	478	10	BF657849	BF657849	ma992f12.
C 254	15.8	79.0	787	14	CA306138	CA306138	C 327	15.4	77.0	478	14	CA555095	CA555095	CO893G09-
C 255	15.8	79.0	796	9	AL580875	AL580875	C 328	15.4	77.0	481	13	BY243831	BY243831	BY243831
C 256	15.8	79.0	800	12	BI932916	EST552805	C 329	15.4	77.0	486	14	CF155910	CF155910	B0630P09-
C 257	15.8	79.0	807	29	CG369334	OG0G026TF	C 330	15.4	77.0	486	28	BH661806	BH661806	BOMF128TR
C 258	15.8	79.0	809	28	BH696645	BOMF181TF	C 331	15.4	77.0	490	28	CG325458	CG325458	TEA105_Ba
C 259	15.8	79.0	834	9	AL529131	AL529131	C 332	15.4	77.0	492	28	BH385412	BH385412	AG-ND_-131
C 260	15.8	79.0	851	10	BF692242	602249113	C 333	15.4	77.0	496	10	AW682703	AW682703	EST00976
C 261	15.8	79.0	856	28	BZ447974	BZ447974	C 334	15.4	77.0	510	12	BG524561	BG524561	43--83_Ste
C 262	15.8	79.0	863	28	B20944	TBEJ3-T7.2 T	C 335	15.4	77.0	515	14	CD291038	CD291038	St-Fu338.
C 263	15.8	79.0	865	14	CK291457	EST754171	C 336	15.4	77.0	530	12	BM403834	BM403834	EST758161
C 264	15.8	79.0	872	29	CG369341	OG0G026TF	C 337	15.4	77.0	530	14	CD290266	CD290266	St-Fu538.
C 265	15.8	79.0	881	28	CG388186	PUMHOX4TB	C 338	15.4	77.0	531	29	CG427151	CG427151	0150726-0
C 266	15.8	79.0	903	12	BG622103	BG622103	C 339	15.4	77.0	534	10	BE404894	BE404894	WHE1206_C
C 267	15.8	79.0	913	9	AL520808	AL520808	C 340	15.4	77.0	536	29	CC799458	CC799458	0150559-0
C 268	15.8	79.0	921	12	BM358157	GA_Ea000	C 341	15.4	77.0	542	10	BF466844	BF466844	UI-M-GCOP
C 269	15.8	79.0	935	10	BF181394	601807280	C 342	15.4	77.0	555	12	BG520816	BG520816	ps52c05..y
C 270	15.8	79.0	938	13	BK328627	BK328627	C 343	15.4	77.0	556	10	BF427832	BF427832	d191g08..y
C 271	15.8	79.0	966	13	BK328628	BK328628	C 344	15.4	77.0	565	28	B2173300	B2173300	CH230_-321
C 272	15.8	79.0	979	28	B08657	B08657	C 345	15.4	77.0	569	12	BM813690	BM813690	EST591783
C 273	15.8	79.0	1011	13	BQ706318	BQ706318	C 346	15.4	77.0	574	28	AQ524050	AQ524050	HS_-5232_B
C 274	15.8	79.0	1060	9	AL575835	AL575835	C 347	15.4	77.0	585	10	BE75870	BE75870	MY_-08-A-0
C 275	15.8	79.0	1096	12	BG329874	602429558	C 348	15.4	77.0	593	10	BF730492	BF730492	ma6b4a06.
C 276	15.8	79.0	1101	28	CC206155	CH261-1318	C 349	15.4	77.0	594	14	CD311937	CD311937	St-Fu691.
C 277	15.8	79.0	1125	10	BF790487	602250157	C 350	15.4	77.0	597	12	B414655	B414655	St-Fu6455
C 278	15.8	79.0	1201	13	BK365741	BK365741	C 351	15.4	77.0	600	10	BG800397	BG800397	2102-29_M
C 279	15.8	79.0	1332	29	CG756498	P051-4-A0	C 352	15.4	77.0	600	12	BI989303	BI989303	4033-72_M
C 280	15.8	79.0	1481	28	CC237544	CH261-191	C 353	15.4	77.0	606	28	A2632296	A2632296	1M0486B23
C 281	15.8	79.0	1560	12	BI413018	602989965	C 354	15.4	77.0	612	29	CE687222	CE687222	ligr-g88-
C 282	15.8	79.0	1674	11	AK053226	Mus_muecu	C 355	15.4	77.0	616	12	B416184	B416184	Bj416184
C 283	15.4	77.0	213	28	BZ096958	CH230-142	C 356	15.4	77.0	618	12	Bu614122	Bu614122	Bu614122
C 284	15.4	77.0	238	10	BM063490	BM063490	C 357	15.4	77.0	618	14	CK222692	CK222692	702401793
C 285	15.4	77.0	244	29	CG583573	OS725890	C 358	15.4	77.0	620	14	CB289435	CB289435	V-B-117A0
C 286	15.4	77.0	261	10	BF554085	UI-R-CO-h	C 359	15.4	77.0	625	14	CA921462	CA921462	EST7639180
C 287	15.4	77.0	263	12	BU323443	BU323443	C 360	15.4	77.0	629	14	BM036010	BM036010	1M0419P07
C 288	15.4	77.0	270	10	BA402442	BA402442	C 361	15.4	77.0	631	12	BM494446	BM494446	IPCGBA1_1
C 289	15.4	77.0	273	29	CC596956	CC596956	C 362	15.4	77.0	634	28	CC050144	CC050144	0150532-0
C 290	15.4	77.0	283	9	AA386834	AA386834	C 363	15.4	77.0	636	12	CE628613	CE628613	ligr-g88-
C 291	15.4	77.0	286	12	BM033939	UI-1-CFO-	C 364	15.4	77.0	639	12	BQ415023	BQ415023	HVMEK000
C 292	15.4	77.0	287	10	BI818148	BI818148	C 365	15.4	77.0	639	29	BX226827	BX226827	Danio_rer
C 293	15.4	77.0	300	10	BM264008	BM264008	C 366	15.4	77.0	647	13	B0167199	B0167199	WHR0058_F
C 294	15.4	77.0	304	10	BM266963	BM266963	C 367	15.4	77.0	648	14	CB059023	CB059023	NISC_jx10
C 295	15.4	77.0	336	10	BM238391	BM238391	C 368	15.4	77.0	650	11	AK005183	AK005183	Mus_muecu
C 296	15.4	77.0	351	10	AM470292	AM470292	C 369	15.4	77.0	654	28	CC1892563	CC1892563	02S2027-0

370	15.4	77.0	658	29	CG400372		443	15.2	76.0	263	9	AA506314
371	15.4	77.0	658	29	CG427792	01S0777-0	C 444	15.2	76.0	264	9	AV238655
372	15.4	77.0	661	10	BF006633	EST435131	C 445	15.2	76.0	271	9	AV259730
373	15.4	77.0	663	28	BH592989	BOHSS59TR	C 446	15.2	76.0	277	10	BB010692
374	15.4	77.0	669	12	BH389950	BH389950	C 447	15.2	76.0	286	10	BB282695
375	15.4	77.0	671	28	BH765743	BH765743	C 448	15.2	76.0	290	9	AL640424
376	15.4	77.0	677	9	AA556716	AA556716 571 Loblo	C 449	15.2	76.0	290	13	BX780974
377	15.4	77.0	678	10	BE790846	601582263	C 450	15.2	76.0	293	10	BB005204
378	15.4	77.0	680	10	BB357252	BB357252	C 451	15.2	76.0	293	12	BG657032
379	15.4	77.0	687	28	BZ495354	BZ495354 BONTT35TF	C 452	15.2	76.0	295	14	N788343
380	15.4	77.0	690	14	CF997829	CF997829	C 453	15.2	76.0	300	9	AV219173
381	15.4	77.0	695	14	CD003924	CD003924 VVA019F07	C 454	15.2	76.0	303	10	BB219213
382	15.4	77.0	711	28	BH085732	BH085732 RPCT-24-9	C 455	15.2	76.0	310	10	BB205297
383	15.4	77.0	712	12	BG523196	29-89 Ste	C 456	15.2	76.0	311	28	AZ454440
384	15.4	77.0	735	28	AO345950	RPCT11-12	C 457	15.2	76.0	318	10	AA902881
385	15.4	77.0	765	14	CA323241	UI-M-FX0-	C 458	15.2	76.0	318	10	BB929446
386	15.4	77.0	782	12	BG468157	602509732	C 459	15.2	76.0	318	10	BB929446
387	15.4	77.0	835	28	BH593136	BOHRQ33TF	C 460	15.2	76.0	326	12	BG554416
388	15.4	77.0	838	29	CNS050204	AL175657 Tetradon	C 461	15.2	76.0	331	9	AA650522
389	15.4	77.0	840	29	CNS050502	AL122316 Tetradon	C 462	15.2	76.0	334	10	BB223035
390	15.4	77.0	843	29	CNS20902	AL294264 Tetradon	C 463	15.2	76.0	342	29	CG928415
391	15.4	77.0	843	29	CNS04K1R	CG20902 CH240 368	C 464	15.2	76.0	344	14	CF506812
392	15.4	77.0	846	12	BG719954	BG719954 602691303	C 465	15.2	76.0	354	10	BF225297
393	15.4	77.0	848	14	CA969599	CA969599 CGLX06423	C 466	15.2	76.0	354	29	CE435638
394	15.4	77.0	849	29	CNS047C5	AL06302 Tetradon	C 467	15.2	76.0	358	28	AQ010566
395	15.4	77.0	857	13	BUS58356	BUS58356	C 468	15.2	76.0	363	10	BE655876
396	15.4	77.0	857	29	CC805534	CC805534 AGENCOURT	C 469	15.2	76.0	363	13	BY213050
397	15.4	77.0	862	29	CG579888	CG579888 ZMMBB-C047	C 470	15.2	76.0	366	28	B2592758
398	15.4	77.0	863	29	CG020988	CG020988 ZMMBB-C055	C 471	15.2	76.0	367	10	AA447338
399	15.4	77.0	864	14	CF829949	CF829949 UCRCR01_0	C 472	15.2	76.0	374	14	CD195151
400	15.4	77.0	870	29	CG517833	CG517833 CH240_364	C 473	15.2	76.0	377	13	BY166002
401	15.4	77.0	873	29	CG954789	CG954789 MBEDH02TR	C 474	15.2	76.0	379	10	AA490991
402	15.4	77.0	887	12	BG300625	BG300625 HVSMED001	C 475	15.2	76.0	380	28	AQ049765
403	15.4	77.0	889	13	BX845255	BX845255	C 476	15.2	76.0	387	10	AA542363
404	15.4	77.0	896	14	CB5608743	CB5608743 AGENCOURT	C 477	15.2	76.0	390	13	BY444993
405	15.4	77.0	912	29	CG968743	CG968743 MBELN56TF	C 478	15.2	76.0	392	10	AA697635
406	15.4	77.0	922	13	BO737035	BO737035 AGENCOURT	C 479	15.2	76.0	395	28	AQ547841
407	15.4	77.0	922	13	BY746278	BY746278	C 480	15.2	76.0	399	28	AQ673057
408	15.4	77.0	952	13	BX341013	BX341013	C 481	15.2	76.0	404	14	CB811908
409	15.4	77.0	980	29	CNS0202TF	AL213188 Tetradon	C 482	15.2	76.0	407	10	BE548993
410	15.4	77.0	986	29	CG900580	CG900580 ZMMBB-C050	C 483	15.2	76.0	411	10	AA343233
411	15.4	77.0	1031	12	BI834473	BI834473 603084638	C 484	15.2	76.0	411	13	BY006359
412	15.4	77.0	1201	9	AL513771	AL513771	C 485	15.2	76.0	414	14	CD012333
413	15.4	77.0	1201	9	AL564116	AL564116	C 486	15.2	76.0	420	28	AQ205240
414	15.4	77.0	1348	12	BM047712	603627450	C 487	15.2	76.0	421	28	AAQ89134
415	15.4	77.0	1458	11	AK047738	AK047738 Mus muscu	C 488	15.2	76.0	422	14	CF372888
416	15.4	77.0	1514	28	CC291468	CC291468 CH261-173	C 489	15.2	76.0	422	14	T90278
417	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 490	15.2	76.0	429	10	AA484023
418	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 491	15.2	76.0	430	29	CC928567
419	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 492	15.2	76.0	431	10	AA479605
420	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 493	15.2	76.0	431	10	BB848002
421	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 494	15.2	76.0	434	14	CD195134
422	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 495	15.2	76.0	436	10	BB849928
423	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 496	15.2	76.0	436	14	CB230886
424	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 497	15.2	76.0	439	9	AI601572
425	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 498	15.2	76.0	439	10	BB626665
426	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 499	15.2	76.0	443	9	AI580295
427	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 500	15.2	76.0	443	9	AI601294
428	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 501	15.2	76.0	445	10	BB726107
429	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 502	15.2	76.0	450	12	BB726107
430	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 503	15.2	76.0	451	9	AI022508
431	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 504	15.2	76.0	451	28	AAQ66857
432	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 505	15.2	76.0	453	10	BF559396
433	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 506	15.2	76.0	454	10	AAW94583
434	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 507	15.2	76.0	455	12	BB816750
435	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 508	15.2	76.0	457	9	AA917696
436	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 509	15.2	76.0	458	10	AAW68757
437	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 510	15.2	76.0	460	28	AAQ80159
438	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 511	15.2	76.0	462	13	BBU78131
439	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 512	15.2	76.0	463	28	BB285560
440	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 513	15.2	76.0	464	9	AAU175638
441	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 514	15.2	76.0	469	9	AAU890531
442	15.4	77.0	1514	28	CC291468	CC291468 Mus muscu	C 515	15.2	76.0	471	12	BBM17916

AA506314 nh45C05.B
AV238655 AV238655
AV259730 AV259730
BB010692 BB010692
BB282695 BB282695
AL640424 AL640424
BX780974 BX780974
BB005204 BB005204
BG657032 BG657032
N788343 N788343
AV219173 AV219173
AT775355 B2756459
BB219213 BB219213
BB205297 BB205297
AZ454440 AZ454440
AA902881 AA902881
BB929446 BB929446
BG554416 BG554416
AA650522 AA650522
BB223035 BB223035
CG928415 CG928415
CF506812 CF506812
BP225297 BP225297
CE435638 CE435638
AQ010566 AQ010566
BE655876 BE655876
BY213050 BY213050
B2592758 B2592758
AA447338 AA447338
CD195151 CD195151
BY166002 BY166002
AA490991 AA490991
AQ049765 AQ049765
AA542363 AA542363
BY444993 BY444993
AA697635 AA697635
AQ547841 AQ547841
AQ673057 AQ673057
CB811908 CB811908
BE548993 BE548993
AA343233 AA343233
BY006359 BY006359
CD012333 CD012333
AQ205240 AQ205240
AAQ89134 AAQ89134
CB849928 CB849928
CB230886 CB230886
AI601572 AI601572
BB626665 BB626665
AI580295 AI580295
AI601294 AI601294
BB726107 BB726107
BB726107 BB726107
BB726107 BB726107
AI022508 AI022508
AAQ66857 AAQ66857
BF559396 BF559396
AAW94583 AAW94583
BB816750 BB816750
AA917696 AA917696
AAW68757 AAW68757
AAQ80159 AAQ80159
BBU78131 BBU78131
BB285560 BB285560
AAU175638 AAU175638
AAU890531 AAU890531
BBM17916 BBM17916

516	15.2	76.0	472	9	A1189543	A1189543 qd22h02.x	C 589	15.2	76.0	574	10	BE354681	BE354681 EST354771
517	15.2	76.0	472	9	A1838722	A1838722 UT-M-A00-	C 590	15.2	76.0	575	28	A2838089	A2838089 ZM0133G08
518	15.2	76.0	472	28	A0660554	A0660554 Sheared D	C 591	15.2	76.0	576	12	BG885427	BG885427 da4d2c08.
519	15.2	76.0	474	10	BE623205	BE623205 ut95a06.x	C 592	15.2	76.0	578	10	AM138767	AM138767 UT-H-B11-
520	15.2	76.0	475	14	CB005274	CB005274 VVC01B03	C 593	15.2	76.0	579	12	BM359158	BM359158 BUJ59158
521	15.2	76.0	477	9	A1935668	A1935668 w099c07.x	C 594	15.2	76.0	579	12	BU359316	BU359316 BUJ59316
522	15.2	76.0	478	12	BI322592	BI322592 kx13h07.y	C 595	15.2	76.0	579	29	CE627130	CE627130 C1gr-g88-
523	15.2	76.0	478	13	BK635970	BK635970 BX635970	C 596	15.2	76.0	580	13	CE627105	CE627105 SJM2BJB08
524	15.2	76.0	479	10	BE114484	BE114484 UT-R-CA0-	C 597	15.2	76.0	581	14	CB916619	CB916619 VVD11G05
525	15.2	76.0	481	28	BE2290	BE2290 RRC11-14H1	C 598	15.2	76.0	583	10	BE694232	BE694232 602082857
526	15.2	76.0	489	9	AA859269	AA859269 UT-R-E0-C	C 599	15.2	76.0	586	12	BM361629	BM361629 BUJ61629
527	15.2	76.0	494	9	A1531862	A1531862 SD03145..5	C 600	15.2	76.0	587	29	CC466857	CC466857 CH240..136
528	15.2	76.0	496	13	BO556161	BO556161 H4038D11-	C 601	15.2	76.0	587	12	BU327803	BU327803 C1gr-g88-
529	15.2	76.0	497	28	AO833959	AO833959 HS 5261.A	C 602	15.2	76.0	588	14	CA892704	CA892704 B0174B03-
530	15.2	76.0	498	12	BU388628	BU388628 BJ388628	C 603	15.2	76.0	591	10	CB899015	CB899015 601285747
531	15.2	76.0	499	14	W20132	W20132 ZB40C10.r1	C 604	15.2	76.0	591	14	CA367755	CA367755 643750.NC
532	15.2	76.0	504	12	BU363161	BU363161 BUJ63161	C 605	15.2	76.0	591	14	CA404168	CA404168 EL01N0513
533	15.2	76.0	504	14	CA208521	CA208521 SCACSB111	C 606	15.2	76.0	591	14	CB425375	CB425375 600352.MA
534	15.2	76.0	506	28	AO243748	AO243748 HS 2057.B	C 607	15.2	76.0	592	12	BU389256	BU389256 BUJ89256
535	15.2	76.0	507	9	A1657550	A1657550 fc15a02.Y	C 608	15.2	76.0	593	29	CC743217	CC743217 ZMMBBD011
536	15.2	76.0	509	9	AA831831	AA831831 oc85g03..8	C 609	15.2	76.0	594	14	CB910032	CB910032 VVD173B01
537	15.2	76.0	509	28	AO776625	AO776625 HS_2148.B	C 610	15.2	76.0	595	14	CA387084	CA387084 668874.NC
538	15.2	76.0	510	14	CB424865	CB424865 599499.MA	C 611	15.2	76.0	596	13	BU357680	BU357680 603476801
539	15.2	76.0	510	14	CF403237	CF403237 CSECS003F	C 612	15.2	76.0	597	29	CE795949	CE795949 C1gr-g88-
540	15.2	76.0	511	12	BU363045	BU363045 BJ363045	C 613	15.2	76.0	600	28	BE708259	BE708259 OGBAP06TC
541	15.2	76.0	519	28	AO505167	AO505167 RPCI-11-2	C 614	15.2	76.0	601	9	AI062238	AI062238 GH01378..5
542	15.2	76.0	520	13	BU809199	BU809199 UL54PH12	C 615	15.2	76.0	601	29	DR46198	ALU737008 Datio_rer
543	15.2	76.0	520	14	CA545244	CA545244 K011E0D5-	C 616	15.2	76.0	602	14	CF486121	CF486121 POL1_35-G
544	15.2	76.0	523	14	CB715589	CB715589 AMGNNUC:N	C 617	15.2	76.0	607	12	BE467408	BE467408 EST12_Pac
545	15.2	76.0	520	28	BH117649	BH117649 RPCI-24-3	C 618	15.2	76.0	607	14	CD646919	CD646919 AUF_106.I
546	15.2	76.0	524	12	BM256385	BM256385 518740.MA	C 619	15.2	76.0	609	9	AI062055	AI062055 GH01103..5
547	15.2	76.0	525	14	CB720171	CB720171 AMGNNUC-U	C 620	15.2	76.0	609	9	AJ397505	AJ397505 AJ397505
548	15.2	76.0	525	28	AO940449	AO940449 Sheared.D	C 621	15.2	76.0	609	13	BU818793	BU818793 UA47BEP05
549	15.2	76.0	526	28	CC064938	CC064938 Igm04004F	C 622	15.2	76.0	609	13	BE267334	BE267334 BA267334
550	15.2	76.0	529	10	AM690744	AM690744 NF03A047S	C 623	15.2	76.0	610	28	AO955222	AO955222 LBRAB06TF
551	15.2	76.0	530	28	AO554837	AO554837 RPCI-11-3	C 624	15.2	76.0	615	9	AM034296	AM034296 EST277967
552	15.2	76.0	534	14	CA545539	CA545539 K0115F12-	C 625	15.2	76.0	615	12	BE791233	BE791233 ESTFNL096
553	15.2	76.0	535	29	CG925610	CG925610 MBENJ23TR	C 626	15.2	76.0	617	13	BE481654	BE481654 DKFED686C
554	15.2	76.0	536	14	CA741071	CA741071 wem1C.pko	C 627	15.2	76.0	617	28	AZ412192	AZ412192 1M0185X01
555	15.2	76.0	537	10	BE749525	BE749525 200496.MA	C 628	15.2	76.0	619	28	A2423496	A2423496 1M0202C09
556	15.2	76.0	540	28	BH015619	BH015619 TDMGC86TH	C 629	15.2	76.0	621	12	BG577046	BG577046 602599461
557	15.2	76.0	542	10	AM237816	AM237816 xm81h09.x	C 630	15.2	76.0	621	14	CA394142	CA394142 C847905.y
558	15.2	76.0	542	10	BE015099	BE015099 127277.MA	C 631	15.2	76.0	621	28	AZ298192	AZ298192 RPCI-23-1
559	15.2	76.0	546	12	BI501358	BI501358 kx31c03.Y	C 632	15.2	76.0	625	13	BU359964	BU359964 603476969
560	15.2	76.0	546	12	BM068107	BM068107 KS08017AL	C 633	15.2	76.0	627	12	BU325128	BU325128 BUJ325128
561	15.2	76.0	547	9	AM792048	AM792048 D00974-R	C 634	15.2	76.0	630	28	AO642252	AO642252 RPCI93-IDP
562	15.2	76.0	548	10	A1813671	A1813671 wJ65d10..x	C 635	15.2	76.0	630	10	AM646759	AM646759 EST327213
563	15.2	76.0	549	14	CD854887	CD854887 DHOAM36Z	C 636	15.2	76.0	632	29	CE496250	CE496250 C1gr-g88-
564	15.2	76.0	549	28	AO992197	AO992197 nbeD0085A	C 637	15.2	76.0	632	12	BG378240	BG378240 UT-R-CV1-
565	15.2	76.0	550	9	A1934581	A1934581 wp09d01.x	C 638	15.2	76.0	633	13	BU402588	BU402588 604140873
566	15.2	76.0	550	12	BI131118	BI131118 G115P631	C 639	15.2	76.0	633	28	BE372491	BE372491 i668B04..b
567	15.2	76.0	550	14	CD647519	CD647519 AUF_108.D	C 640	15.2	76.0	634	28	BZ881424	BZ881424 CH240..266
568	15.2	76.0	551	12	BI386293	BI386293 BFL26..000	C 641	15.2	76.0	638	13	BE267333	BE267333 BX267333
569	15.2	76.0	551	12	BU359641	BU359641 BJ359641	C 642	15.2	76.0	638	14	CB321897	CB321897 UT-CF-FNO
570	15.2	76.0	553	10	AME57940	AME57940 L0288B11-	C 643	15.2	76.0	639	28	BE633624	BE633624 SALK..0428
571	15.2	76.0	554	14	CF162722	CF162722 B0715H02-	C 644	15.2	76.0	640	10	BE021763	BE021763 BR021763
572	15.2	76.0	555	14	CB613545	CB613545 AMGNNUC-U	C 645	15.2	76.0	640	14	CD190714	CD190714 WS1..0064U
573	15.2	76.0	555	28	BZ708270	BZ708270 OGBAP06TM	C 646	15.2	76.0	640	28	BE292525	BE292525 CH240..72E
574	15.2	76.0	557	29	CG971340	CG971340 MBEPF08TF	C 647	15.2	76.0	643	12	BI386954	BI386954 BFL26..001
575	15.2	76.0	560	13	BU948965	BU948965 in71e03.Y	C 648	15.2	76.0	643	12	BI649022	BI649022 603277202
576	15.2	76.0	561	28	AZ054783	AZ054783 RPCI-23-4	C 649	15.2	76.0	646	28	BZ805871	BZ805871 PUGAL54TB
577	15.2	76.0	563	9	A1658278	A1658278 fc20h09.Y	C 650	15.2	76.0	647	28	AZ804958	AZ804958 2M0066D12
578	15.2	76.0	564	10	AM147848	AM147848 dai0e06.x	C 651	15.2	76.0	647	29	BE113495	BE113495 Datio_rer
579	15.2	76.0	564	14	CD731776	CD731776 4041736.I	C 652	15.2	76.0	650	13	BX667941	BX667941 BX667941
580	15.2	76.0	566	12	BI964009	BI964009 i6e3g04.Y	C 653	15.2	76.0	650	28	BE2448956	BE2448956 BONGM57TR
581	15.2	76.0	566	14	CB098319	CB098319 iue3b01.Y	C 654	15.2	76.0	652	10	BE056531	BE056531 894010A05
582	15.2	76.0	569	10	BE475053	BE475053 BP71F10.Y	C 655	15.2	76.0	653	14	CA501293	CA501293 WHR54031.G
583	15.2	76.0	569	13	BM291970	BM291970 BM291970	C 656	15.2	76.0	656	10	BE670697	BE670697 7e33h07..x
584	15.2	76.0	569	29	CB828021	CB828021 C1gr-g88-	C 657	15.2	76.0	658	13	BT738487	BT738487 BT738487
585	15.2	76.0	571	14	CA825254	CA825254 R56A08.tw	C 658	15.2	76.0	659	10	AM612120	AM612120 hg94d08.x
586	15.2	76.0	572	29	CE694107	CE694107 C1gr-g88-	C 659	15.2	76.0	660	13	BO240611	BO240611 TAB05015B
587	15.2	76.0	573	10	AM474485	AM474485 xs97d01.x	C 660	15.2	76.0	661	29	CG219902	CG219902 CG4AL10TC
588	15.2	76.0	573	29	AG331466	AG331466 Lotus cor	C 661	15.2	76.0	662	14	CD719814	CD719814 VBA165D11

662	15.2	76.0	666	28	BH265692	CH230-66G	c 735	15.2	76.0	756	13	BU314869	BU314869	603854857
c 663	15.2	76.0	667	13	BU664531	C1118C11.	736	15.2	76.0	757	28	BI6688390	BI6688390	603316309
664	15.2	76.0	667	14	CD715115	VVB820BD02	737	15.2	76.0	757	12	CC429670	CC429670	PUBEL537
c 665	15.2	76.0	668	12	BZ448957	BONSB47TF	c 738	15.2	76.0	758	13	BO794261	BO794261	EST 3129
c 666	15.2	76.0	668	12	BM104581	fx49a07.y	c 739	15.2	76.0	759	14	CB528529	CB528529	UT-H-TF2-
c 667	15.2	76.0	669	9	AL711567	DFFZ6861	c 740	15.2	76.0	761	28	CC438445	CC438445	PUBH933TD
c 668	15.2	76.0	669	29	CC548242	CH240_432	741	15.2	76.0	769	14	CD648590	CD648590	AUF_103_ID
c 669	15.2	76.0	669	29	CC316817	OGMT53TH	742	15.2	76.0	770	29	CC136646	CC136646	NDI..83E22
c 670	15.2	76.0	672	14	BY753784	BY753784	743	15.2	76.0	771	28	CG855466	CG855466	ZMMBCC022
c 671	15.2	76.0	674	29	CG179095	PUI7A83TB	c 744	15.2	76.0	771	29	CG855297	CG855297	ZMMBCC022
c 672	15.2	76.0	676	14	CP913281	A0646B07	c 745	15.2	76.0	774	29	CG855297	CG855297	ZMMBCC022
c 673	15.2	76.0	678	28	BZ659619	OGAMS69TM	746	15.2	76.0	776	14	CB328506	CB328506	UT-R-F50-
c 674	15.2	76.0	678	12	BI955051	HVSMEM002	c 747	15.2	76.0	777	12	BI666180	BI666180	603287176
c 675	15.2	76.0	679	12	BI310778	EST531252	748	15.2	76.0	777	14	CF518202	CF518202	CAPE0006_I
c 676	15.2	76.0	680	28	BH322453	CH230-104	749	15.2	76.0	782	28	BZ757169	BZ757169	PURBZ85TD
c 677	15.2	76.0	680	29	CG921533	MBEHR88TF	c 750	15.2	76.0	782	29	CG681188	CG681188	ZMMBCC014
c 678	15.2	76.0	681	13	BU417073	603670803	c 751	15.2	76.0	785	28	BH026368	BH026368	RPCT-24-3
c 679	15.2	76.0	681	29	CE028127	L19T-g98-	c 752	15.2	76.0	786	12	BM410861	BM410861	UT-M-EV0-
c 680	15.2	76.0	683	13	BU621226	UI-H-FL1-	c 753	15.2	76.0	787	13	BU364970	BU364970	602447518
c 681	15.2	76.0	684	13	BK275836	BK275836	c 754	15.2	76.0	788	12	BG424583	BG424583	602447518
c 682	15.2	76.0	686	12	BI931372	EST551261	c 755	15.2	76.0	788	13	BU377525	BU377525	603812676
c 683	15.2	76.0	686	28	BZ659608	OGAMS69TC	c 756	15.2	76.0	789	13	BO442870	BO442870	UT-M-EV0-
c 684	15.2	76.0	692	14	CA852105	E03C01_C0	c 757	15.2	76.0	793	11	AK006524	AK006524	Mus muscu
c 685	15.2	76.0	693	12	BM971539	UI-CF-ECL	c 758	15.2	76.0	793	29	CNS07265	CNS07265	clone BAO
c 686	15.2	76.0	693	14	CB975340	CAB30006_	c 759	15.2	76.0	796	13	BU972947	BU972947	CAB30002
c 687	15.2	76.0	697	13	BK845767	BK845767	c 760	15.2	76.0	798	14	CG692000	CG692000	ZMMBCC013
c 688	15.2	76.0	697	29	ACG09884	Pan troG1	c 761	15.2	76.0	798	29	CG500000	CG500000	PUBFX54TD
c 689	15.2	76.0	698	12	BP009613	BP009613	c 762	15.2	76.0	799	29	CG104284	CG104284	CBUFD0004
c 690	15.2	76.0	699	28	BZ805874	PUGAL54TD	c 763	15.2	76.0	800	14	CF512668	CF512668	clone BAO
c 691	15.2	76.0	701	13	BY750649	BY750649	c 764	15.2	76.0	801	29	CC849320	CC849320	NDL..30M14
c 692	15.2	76.0	702	9	AJ395348	AJ395348	c 765	15.2	76.0	803	29	CG872822	CG872822	ZMMBCC021
c 693	15.2	76.0	704	29	CG696016	pastrbac00	c 766	15.2	76.0	804	28	CC076830	CC076830	clone K33r.
c 694	15.2	76.0	705	10	BE273085	601142721	c 767	15.2	76.0	804	29	CNS072M3	CNS072M3	clone BAO
c 695	15.2	76.0	705	12	BM018151	603645747	c 768	15.2	76.0	812	14	CD648583	CD648583	AUF_103_C
c 696	15.2	76.0	708	9	AV704562	AV704562	c 769	15.2	76.0	813	14	CD764178	CD764178	PUBDC69TD
c 697	15.2	76.0	710	28	AZ045328	nbeeb0051D	c 770	15.2	76.0	813	28	BZ721719	BZ721719	PUBDC69TD
c 698	15.2	76.0	711	13	BU458089	603774375	c 771	15.2	76.0	814	13	BU333583	BU333583	603498545
c 699	15.2	76.0	712	28	AQ272004	nbxb0027G	c 772	15.2	76.0	814	14	CG450985	CG450985	CAB20004
c 700	15.2	76.0	714	12	BI891088	ZF637-3-0	c 773	15.2	76.0	814	29	CG450985	CG450985	ZMMBCC002
c 701	15.2	76.0	715	10	BF020211	uw62e02.x	c 774	15.2	76.0	815	12	BO265558	BO265558	clone BAO
c 702	15.2	76.0	715	14	CA419738	UI-H-FH0-	c 775	15.2	76.0	815	14	CB559013	CB559013	AGENCOCURT
c 703	15.2	76.0	716	14	CA500469	WHE4020_A	c 776	15.2	76.0	815	28	CC366032	CC366032	PUBCM48TB
c 704	15.2	76.0	719	28	BH954470	od166c10.	c 777	15.2	76.0	818	10	BE622141	BE622141	601440703
c 705	15.2	76.0	723	14	CB940680	IPCCG14_	c 778	15.2	76.0	820	10	BE515008	BE515008	601316870
c 706	15.2	76.0	724	14	CD405274	Gm CK2854	c 779	15.2	76.0	820	28	BH602949	BH602949	BOGRQ73TR
c 707	15.2	76.0	724	14	CF520295	AGENCOCURT	c 780	15.2	76.0	827	13	BU333165	BU333165	603542864
c 708	15.2	76.0	724	28	BH088729	RPCT-24-3	c 781	15.2	76.0	828	12	BM159808	BM159808	EST562331
c 709	15.2	76.0	724	28	BH091330	CH230-193	c 782	15.2	76.0	828	13	BU343004	BU343004	603518921
c 710	15.2	76.0	725	12	BI933375	EST553264	c 783	15.2	76.0	828	28	BZ612016	BZ612016	WHADEP7TR
c 711	15.2	76.0	727	14	CD650410	CVG110111	c 784	15.2	76.0	830	14	CB999478	CB999478	AGENCOCURT
c 712	15.2	76.0	732	12	BI930980	EST550869	c 785	15.2	76.0	833	13	BX771746	BX771746	603542864
c 713	15.2	76.0	736	13	BO604795	MI-P-CP1-	c 786	15.2	76.0	833	29	CG269729	CG269729	OGODH90TV
c 714	15.2	76.0	737	13	BO604795	MI-P-CP1-	c 787	15.2	76.0	834	14	CK092088	CK092088	G065P17_3
c 715	15.2	76.0	739	13	BU245341	603782136	c 788	15.2	76.0	835	12	BI697829	BI697829	603343943
c 716	15.2	76.0	740	14	CF347937	AGENCOCURT	c 789	15.2	76.0	838	28	CC444727	CC444727	PURHRT2TB
c 717	15.2	76.0	740	29	CG999064	ZMMBB052	c 790	15.2	76.0	838	29	CNS07467	CNS07467	clone BAO
c 718	15.2	76.0	741	10	BF686195	963102F09	c 791	15.2	76.0	842	13	AG904894	AG904894	ZMMBCC012
c 719	15.2	76.0	742	14	CF536641	UI-M-G10-	c 792	15.2	76.0	843	29	CG016722	CG016722	AGENCOCURT
c 720	15.2	76.0	742	28	AZ087043	RPCT-23-6	c 793	15.2	76.0	844	12	BI257143	BI257143	602973164
c 721	15.2	76.0	744	12	BI931321	EST551210	c 794	15.2	76.0	845	13	BU561126	BU561126	AGENCOCURT
c 722	15.2	76.0	744	13	BU112932	603130703	c 795	15.2	76.0	847	28	CC104788	CC104788	CSU-K34.1
c 723	15.2	76.0	745	29	CG688644	ZMMBCC012	c 796	15.2	76.0	848	28	CG834601	CG834601	ZMMBCC014
c 724	15.2	76.0	746	13	BO509810	EST617225	c 797	15.2	76.0	850	28	BZ997900	BZ997900	PUBDZ88TD
c 725	15.2	76.0	747	14	CA227212	SCVPEL104	c 798	15.2	76.0	856	14	CK155565	CK155565	FGAS03639
c 726	15.2	76.0	748	14	CP182443	UI-M-EV0-	c 799	15.2	76.0	860	13	BU374082	BU374082	603587731
c 727	15.2	76.0	749	14	CB973007	CAB30002_	c 800	15.2	76.0	864	13	BU362470	BU362470	603786694
c 728	15.2	76.0	749	28	AQ269606	HS_2036_E								
c 729	15.2	76.0	750	13	BU030742	603739472								
c 730	15.2	76.0	750	28	BZ843734	CH240_205								
c 731	15.2	76.0	751	10	BZ815375	BB615375								
c 732	15.2	76.0	753	28	BH317037	CH230-120								
c 733	15.2	76.0	755	29	CC921953	LO60104B								
c 734	15.2	76.0	756	10	BE390344	601285372								

RESULT 1
BG201597/c

ALIGNMENTS

LOCUS BG201597 219 bp mRNA linear EST 21-APR-2001
 DEFINITION RST0938 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG201597
 VERSION BG201597.1 GI:137323284
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 219)
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
 Creation of genome-wide protein expression libraries using random activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL MEDLINE PUBLISHED 21227151
 COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@athersys.com
 High quality sequence stop: 219.
 Location/Qualifiers
 1..219
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

FEATURES
 source

ORIGIN
 Query Match 100.0%; Score 20; DB 12; Length 219;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GGTCCTACTCATGTTGATG 20
 ||||||||||||||||
 72 GGTCCTACTCATGTTGATG 53

Db

RESULT 2
 LOCUS BG212172 221 bp mRNA linear EST 21-APR-2001
 DEFINITION RST1643 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG212172
 VERSION BG212172.1 GI:13733747
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 221)
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
 Creation of genome-wide protein expression libraries using random activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL MEDLINE PUBLISHED 21227151
 COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@athersys.com
 High quality sequence stop: 221.
 Location/Qualifiers
 1..221
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

FEATURES
 source

COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@athersys.com
 High quality sequence stop: 221.
 Location/Qualifiers
 1..221
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

FEATURES
 source

ORIGIN
 Query Match 100.0%; Score 20; DB 12; Length 221;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GGTCCTACTCATGTTGATG 20
 ||||||||||||||||
 74 GGTCCTACTCATGTTGATG 55

Db

RESULT 3
 LOCUS BG185555 352 bp mRNA linear EST 21-APR-2001
 DEFINITION RST4504 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG185555
 VERSION BG185555.1 GI:13707242
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 352)
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
 Creation of genome-wide protein expression libraries using random activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL MEDLINE PUBLISHED 11329013
 COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@athersys.com
 High quality sequence stop: 352.
 Location/Qualifiers
 1..352
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

FEATURES
 source

REFERENCE 1 (bases 1 to 408)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Huang,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z., and Han,Z.
TITLE Insight into hepatocellular carcinoma at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..408
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKCDCA01"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKC"
/note="Vector: pBlueScript sk(-); Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 408;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGTCCATCTCATGTTGATG 20
|||||
120 GGTCCATCTCATGTTGATG 101

RESULT 7
AV690967/c 408 bp mRNA linear EST 16-JAN-2002
DEFINITION AV690967 GKC Homo sapiens cDNA clone GKCD005 5', mRNA sequence.
ACCESSION AV690967
VERSION AV690967.1 GI:10292830
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 408)
Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Huang,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z., and Han,Z.
TITLE Insight into hepatocellular carcinoma at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers

source
1..408
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKCD005"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKC"
/note="Vector: pBlueScript sk(-); Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 408;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGTCCATCTCATGTTGATG 20
|||||
120 GGTCCATCTCATGTTGATG 101

RESULT 8
A1961077 437 bp mRNA linear EST 08-MAR-2000
LOCUS A1961077
DEFINITION wq44901.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2474160 3' similar to gb:D14531 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS (HUMAN);, mRNA sequence.
ACCESSION A1961077
VERSION A1961077.1 GI:5753858
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 437)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNt at: www-bio.ln1l.gov/bbtp/image/image.html
Insert Length: 766 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source
1..437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2474160"
/tissue_type="pooled germ cell tumors"
/lab_host="HDH08"
/clone_lib="NCI CGAP GC6"
/note="Vector: pT7T5D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP GC4 was prepared, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneds1257096-1258631, 1469064-1470963, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Location/Qualifiers

Query Match 100.0%; Score 20; DB 9; Length 437;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
 158 GGTCCATCTCATGTTGATG 177

Db

RESULT 9
 BU679004 461 bp mRNA linear EST 07-OCT-2002
 LOCUS UI-CF-DUI-aat-o-05-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone
 DEFINITION UI-CF-DUI-aat-o-05-0-UI 3', mRNA sequence.
 BU679004
 VERSION BU679004.1 GI:23526515
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 461)
 REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.
 AUTHORS Normalization and subtraction: two approaches to facilitate gene
 TITLE discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.reagen.com) or from Open Biosystems
 (www.openbiosystems.com).
 The following repetitive elements were found in this cDNA
 sequence: 49-75, >AT rich#low_complexity (matched complement)
 Seq primer: M13 FORWARD
 POLY-A=yes.

FEATURES
 source Location/Qualifiers
 1..461
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-DUI-aat-o-05-0-UI"
 /tissue_type="Primary Lung Epithelial Cells"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-DUI"
 /note="Organ: Lung; Vector: pTV73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I site_2: Not I;
 UI-CF-DUI is a normalized cDNA library containing the
 following tissue(s): Primary Lung Epithelial Cells The
 library was constructed according to Bonaldo, Lennon and
 Soares, Genome Research, 6:791-806, 1996. First strand
 cDNA synthesis was primed with an oligo-dT primer
 containing a Not I site. Double stranded cDNA was ligated
 to an EcoR I adaptor, digested with Not I, and cloned
 directionally into pTV73-Pac vector. The oligonucleotide
 used to prime the synthesis of first-strand cDNA contains
 a library tag sequence that is located between the Not I
 site and the (dT)18 tail. The sequence tag for this
 library is GGCTGTAGC.
 TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
 TAG_LIB=UI-CF-DUI
 TAG_SEQ=GGCTGTAGC"

ORIGIN
 Query Match 100.0%; Score 20; DB 13; Length 461;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
 175 GGTCCATCTCATGTTGATG 194

Db

RESULT 10
 AA828925 486 bp mRNA linear EST 07-APR-1998
 LOCUS o881a10.s1 NCI CGAP Ov2 Homo sapiens cDNA clone IMAGE:1374330
 DEFINITION similar to gp:DJ4533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS
 (HUMAN);, mRNA sequence.
 AA828925
 VERSION AA828925.1 GI:2902024
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 486)
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strusberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILMIL at:
 www.bio.lnlnl.gov/bbrp/image/image.html
 Insert length: 528 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 353.

FEATURES
 source Location/Qualifiers
 1..486
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1374330"
 /sex="female"
 /tissue_type="ovary"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Ov2"
 /note="Vector: pAMP10; mRNA made from invasive ovarian
 tumor, cDNA made by oligo-dT priming. Non-directionally
 cloned. Size-selected on agarose gel, average insert size
 600 bp. Reference: Krizman et al. (1996) Cancer Research
 56:5380-5383."

ORIGIN
 Query Match 100.0%; Score 20; DB 9; Length 486;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
 |||||
 149 GGTCCATCTCATGTTGATG 168

Db

RESULT 11
 BI793276 487 bp mRNA linear EST 12-MAR-2002
 LOCUS i650b09.y1 Melton Normalized Human Islet 4 M4-HIS 1 Homo sapiens
 DEFINITION cDNA clone IMAGE:5670113 5' similar to SW:XP_A_HUMAN P23025

DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS ;, mRNA sequence.
 ACCESSION BI793276
 VERSION BI793276.1 GI:15821001
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 487)
 REFERENCE Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scaer, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistein, A., Schmitt, A., Theising, B., Rutter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other ESTs: 1650D09.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohpc.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu) This sequence now available from the IMAGE consortium. For clone orders contact: info@image.llnl.gov
 Seq primer: -40RP from Gibco
 High quality sequence stop: 261.
 Location/Qualifiers
 1..487
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5670113"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1; Starting library constructed using Superscript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."
 ORIGIN
 Query Match 100.0%; Score 20; DB 12; Length 487;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTTGATG 20
 |||||
 Db 459 GGTCCATCTCATGTTGATG 440
 RESULT 12
 LOCUS AA167098/c 506 bp mRNA linear EST 09-MAR-1998
 DEFINITION zp05g10.r1 Striatogene ovarian cancer (#937219) Homo sapiens cDNA

clone IMAGE:595554 5' similar to gb:D14533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS (HUMAN);, mRNA sequence.
 ACCESSION AA167098
 VERSION AA167098.1 GI:1745474
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 506)
 REFERENCE Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Joer, S., Kirizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepcoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 TITLE
 JOURNAL
 COMMENT
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 667 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 364.
 Location/Qualifiers
 1..506
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:595554"
 /sex="Female"
 /dev_stage="adult, 64 years"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Striatogene ovarian cancer (#937219)"
 /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Papillary serous carcinoma, isolated from ascites, 64 year old caucasian. Average insert size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACACGAC 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"
 ORIGIN
 Query Match 100.0%; Score 20; DB 9; Length 506;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTTGATG 20
 |||||
 Db 246 GGTCCATCTCATGTTGATG 227
 RESULT 13
 LOCUS BQ186624/c 510 bp mRNA linear EST 30-APR-2002
 DEFINITION UI-E-EJ1-ej-r-n-03-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
 ACCESSION BQ186624
 VERSION BQ186624.1 GI:20362175
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 510)
 REFERENCE Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL
 MEDLINE
 PUBMED
 8889548

COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.reagen.com).
The following repetitive elements were found in this CDNA sequence: 432-458, >AT rich#Low_complexity (matched compliment)
Seq primer: M13 REVERSE.

FEATURES
source
1..510
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-n-03-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EJ1 is a subcloned CDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGATTCAGAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN
Query Match 100.0%; Score 20; DB 13; Length 510;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20
|||||
Db 332 GGTCCTACTCATGTTGATG 313

RESULT 14
AA166911 513 bp mRNA linear EST 09-MAR-1998
LOCUS AA166911
DEFINITION zp05g10.s1 Stratagene ovarian cancer (#937219) Homo sapiens CDNA
clone IMAGE:595554 3 similar to gb:D14533 DNA-REPAIR PROTEIN
COMPLEMENTING XP-A CELLS (HUMAN); contains MBR37.ct2 MBR37 repetitive
element ; mRNA sequence.
AA166911
VERSION AA166911.1 GI:1745421
EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 513)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,
Kizman, D., Kucaba, T., Lacy, W., Le, N., Lennon, G., Marra, M.,

TITLE
JOURNAL
COMMENT
Martin, J., Moore, B., Schejlenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 667 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 263.

FEATURES
source
1..513
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:595554"
/sex="female"
/dev_stage="adult, 64 years"
/lab_host="SCOR (kanamycin resistant)"
/clone_lib="Stratagene ovarian cancer (#937219)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Papillary serous carcinoma, isolated from ascites, 64 year old caucasian. Average insert size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'."

ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 513;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20
|||||
Db 357 GGTCCTACTCATGTTGATG 376

RESULT 15
B0184648 515 bp mRNA linear EST 30-APR-2002
LOCUS B0184648
DEFINITION UI-E-EJ1-n-03-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
UI-E-EJ1-n-03-0-UI 3', mRNA sequence.
B0184648
VERSION B0184648.1 GI:20360199
EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 515)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research

Genetics (www.reegen.com).
The following repetitive elements were found in this cDNA
Sequence: 49-75, >AT rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes

FEATURES

source

Location/Qualifiers
1..515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-B-EJ1-ajr-n-03-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-B-EJ1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-B-EJ1 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATTCACAGA; lens, CGATTAGCGA; eye anterior segment,
AATGCCGCGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina
Foveal and Macular, GTCG; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI).
TAG_LIB=UI-B-EJ1
TAG_SEQ=GTCG"

ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 515;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
Db 175 GGTCCATCTCATGTTGATG 194

RESULT 16 534 bp mRNA linear EST 24-MAY-2000
AM890854
LOCUS
DEFINITION R3-NT0053-100500-021-G11 NT0053 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM890854
VERSION AM890854.1 GI:8055059
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 534)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Nagai, M.A., da Silva, W. Jr., Zagzo, M.A., Bordin, S., Costa, F.P.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202653
PUBMED 10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/ICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RCS-NT0053-100
500-021-G1453-2000-05-10x4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 486.

FEATURES

source

Location/Qualifiers
1..534
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0053"
/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 534;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
Db 430 GGTCCATCTCATGTTGATG 449

RESULT 17 587 bp mRNA linear EST 17-JAN-2003
CB048344
LOCUS
DEFINITION NISC_g104c05.y1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3270537
5', mRNA sequence.
ACCESSION CB048344
VERSION CB048344.1 GI:27786631
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 587)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nlm.nih.gov
CDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
info@image.lnl.gov
Plate: LLM8006 row: F column: 10
Seq primer: M13Rpl reverse primer (ABI).
Location/Qualifiers
1..587
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3270537"
/sex="male"

/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 587;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
214 GGTCCATCTCATGTTGATG 195

Db

RESULT 18
CB048343 591 bp mRNA linear EST 17-JUN-2003
LOCUS NISC g104c05.x1 NCI CGAP Pr28 Homo sapiens CDNA IMAGE:3270537
DEFINITION 3', mRNA sequence.
ACCESSION CB048343
VERSION CB048343.1 GI:27786630
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 591)
NCI-CCAP htcp://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgepbbs-remail.nih.gov
CDNA Library Preparation:
DNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: L1AM8006 row: F column: 10
Seq primer: -21m13 forward primer (ABI).
Location/Qualifiers
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3270537"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 591;

Best Local Similarity 100.0%; Pred. No. 85;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
378 GGTCCATCTCATGTTGATG 397

Db

RESULT 19
AW979099 592 bp mRNA linear EST 02-JUN-2000
LOCUS EST931209 MAGE resequences, MAGP Homo sapiens CDNA, mRNA sequence.
DEFINITION AW979099
ACCESSION AW979099.1 GI:8170384
VERSION AW979099.1 GI:8170384
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 592)
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeaman, T. J. and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: john@tigr.org
Plate: 402
Seq primer: Forward.
Location/Qualifiers
1..592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGP"
/note="Vector: pBluescriptsm"

ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 592;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
270 GGTCCATCTCATGTTGATG 289

Db

RESULT 20
BM739320 595 bp mRNA linear EST 01-MAR-2002
LOCUS K-EST0008760 S2SNUG68 Homo sapiens CDNA clone S2SNUG68-6-H08 5',
DEFINITION mRNA sequence.
ACCESSION BM739320
VERSION BM739320.1 GI:19060649
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 595)
Kim, N. S., Hahn, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Chu, M. Y., Kim, M. R.,
Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and
Kim, Y. S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology

52 Eooun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 6 row: H column: 08
High quality sequence stop: 595.
Location/Qualifiers

FEATURES

source

```
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S2SN0668-6-H08"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-668"
/lab_host="Top10F"
/clone_lib="S2SN0668"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The cDNA vector
vector was adjusted to have about 60nt. The dt-tailed
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F by electroporation
method."
```

ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 595;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
DB 412 GGTCCATCTCATGTTGATG 393

RESULT 21
BI792983 598 bp mRNA linear EST 12-MAR-2002
LOCUS i650b09.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
DEFINITION cDNA clone IMAGE:5670113 similar to SW:XP4.HUMAN P23025
DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS ; mRNA sequence.
BI792983
ACCESSION BI792983.1 GI:15820708
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 598)
REFERENCE Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
AUTHORS Lemstra, L., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blisstein, A.,
Schiller, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Taagereishvili, R.,
Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
Seq primer: -40UP from Gldco
High quality sequence stop: 451.
Location/Qualifiers

FEATURES

source

```
1..598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5670113"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not I;
Site 2: Sal I; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dt priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."
```

ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 598;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
|||||
DB 354 GGTCCATCTCATGTTGATG 373

RESULT 22
BF446397 627 bp mRNA linear EST 01-DEC-2000
LOCUS 7p36g05.x1 NCI CGAP P128 Homo sapiens cDNA clone IMAGE:3648176 3'
DEFINITION similar to SW:XP4.HUMAN P23025 DNA-REPAIR PROTEIN COMPLEMENTING
XP-A CELLS ;contains MER37.t2 MER37 repetitive element ; mRNA
sequence.
BF446397
ACCESSION BF446397.1 GI:11511535
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 627)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gldco
High quality sequence stop: 460.

```

FEATURES
  source
    Location/Qualifiers
      1..627
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:3648176"
        /sex="male"
        /dev_stage="adult"
        /lab_host="DH10B"
        /clone_11b="NCI_CGAP_Pr28"
        /note="Organ: Prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match      100.0%; Score 20; DB 10; Length 627;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20
    |||||
Db 353 GGTCCATCTCATGTTGATG 372

RESULT 23
LOCUS BM853017/c 638 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0134360 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-131-D04
5', mRNA sequence.
ACCESSION BM853017
VERSION BM853017.1 GI:19209416
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 638)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 131 row: D column: 04
High quality sequence stop: 638.
Location/Qualifiers
  1..638
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="S22SNU16n1-131-D04"
    /sex="F"
    /tissue_type="ascites"
    /cell_type="Lymphoblast-like"
    /cell_line="SNU-16"
    /lab_host="DH10B"
    /clone_11b="S22SNU16n1"
    /note="Organ: Stomach; Vector: pT7T3-Pac; Site_1: EcoRI;
Site_2: NotI; The S22SNU16 library was contributed by the
Soares laboratory and it was constructed as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome

```

```

Research 6(9): 791-806. RNA was prepared from harvested
cells of SNU-16 culture. SNU-16 cell was obtained from
Korean Cell Line Bank (KCLB). SNU-16 was established from
ascitic fluids of Korean patients by Park J.G. et al.
(1990), Cancer Res 50: 2773-2780."

ORIGIN

Query Match      100.0%; Score 20; DB 12; Length 639;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20
    |||||
Db 69 GGTCCATCTCATGTTGATG 50

RESULT 24
LOCUS BI911749/c 639 bp mRNA linear EST 16-OCT-2001
DEFINITION 603065343p1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214347 5',
mRNA sequence.
ACCESSION BI911749
VERSION BI911749.1 GI:16175618
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 639)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM11538 row: f column: 12
High quality sequence start: 3
High quality sequence stop: 632.
Location/Qualifiers
  1..639
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5214347"
    /tissue_type="leukocyte"
    /lab_host="DH10B"
    /clone_11b="NIH_MGC_118"
    /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb. Insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match      100.0%; Score 20; DB 12; Length 639;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20
    |||||
Db 619 GGTCCATCTCATGTTGATG 600

```

RESULT 25
LOCUS CB305774 672 bp mRNA linear EST 04-MAR-2003
DEFINITION UI-CF-EN1-aeg-k-02-0-UI.s1 UI-CF-EN1 Homo sapiens CDNA clone
ACCESSION CB305774
VERSION CB305774.1 GI:28846285
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 672)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtrraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul.mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.reagen.com) or from Open Biosystems (www.openbiosystems.com).
The following repetitive elements were found in this cDNA sequence: 495-521, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source
1..672
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-aeg-k-02-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGCT.
TAG_TISSUE=Human Lung Epithelial Cell lines untreated LPS 6hr to LPS 24h
TAG_LIB=UI-CF-EN1
TAG_SEQ=CTGCTCAGCT"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 672;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGTGATG 20

Db 621 GGTCATCTCATGTGTGATG 640
|||||
RESULT 26
LOCUS BM992090 678 bp mRNA linear EST 17-JUN-2002
DEFINITION UI-H-DP1-auf-m-16-0-UI.s1 NCI CGAP_DP1 Homo sapiens CDNA clone
ACCESSION BM992090
VERSION BM992090.1 GI:19711479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 678)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA sequence: 501-527, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source
1..678
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5868855"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DP1"
/note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_DP1 is a normalized cDNA library containing the following tissue(s): Subchondral Bone. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTTAGCGTC.
TAG_TISSUE=Subchondral bone
TAG_LIB=UI-H-DP1
TAG_SEQ=GTTAGCGTC"

ORIGIN
Query Match 100.0%; Score 20; DB 12; Length 678;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGTGATG 20
|||||
Db 627 GGTCATCTCATGTGTGATG 646

RESULT 27
CA450156

LOCUS CA450156 705 bp mRNA linear EST 08-NOV-2002
DEFINITION UI-CF-FNO-aff-k-13-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
ACCESSION CA450156
VERSION CA450156.1 GI:24814576
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bonyfishes; Actinopterygii; Clupeiformes; Clupeidae; Homo.
REFERENCE 1 (bases 1 to 705)
AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 89704477
PubMed 8895548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul.mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 499-525, >AT rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLY(A)=Yes.
FEATURES
source
Location/Qualifiers
1..705
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-aff-k-13-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pTZ19-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-FNO is a subtracted cDNA library derived from two
normalized Human Lung epithelial cell libraries (EN1 and
DU1) The library was subtracted according to Bernaldo,
Lennon and Soares, Genome Research, 6:791-806,
1996. For additional information, contact:
bento-soares@uiowa.edu
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_LIB=UI-CF-FNO
TAG_SEQ=GGCTGATGCGC"
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 705;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTCATCTCATGTTGATG 20
|||||
DB 625 GGTCATCTCATGTTGATG 644
RESULT 28
LOCUS B0772342 748 bp mRNA linear EST 26-JUL-2002
DEFINITION UI-H-EZ1-bhl-1-20-0-UI.s1 NCI CGAP Ch2 Homo sapiens cDNA clone
ACCESSION B0772342
VERSION B0772342.1 GI:21980818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bonyfishes; Actinopterygii; Clupeiformes; Clupeidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS RST5654 Athersys RACE Library
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strassberg, Ph.D.
Email: cga@bcr-remail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of
Orthopedics
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 499-525, >AT rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLY(A)=Yes.
FEATURES
source
Location/Qualifiers
1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EZ1-bhl-1-20-0-UI"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Pelvis; Vector: pTZ19-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_Ch2 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma Grade II. The library
was constructed according to Bernaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTZ19-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TGATCAGCGT.
TAG_TISSUE=grade-2-chondrosarcoma
TAG_LIB=UI-H-EZ1
TAG_SEQ=ATCTAATATG"
ORIGIN
Query Match 100.0%; Score 20; DB 13; Length 748;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTCATCTCATGTTGATG 20
|||||
DB 625 GGTCATCTCATGTTGATG 644
RESULT 29
LOCUS BG186679/c 790 bp mRNA linear EST 21-APR-2001
DEFINITION RST5654 Athersys RACE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG186679
VERSION BG186679.1 GI:13708366
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bonyfishes; Actinopterygii; Clupeiformes; Clupeidae; Homo.
REFERENCE 1 (bases 1 to 790)

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bonyfishes; Actinopterygii; Clupeiformes; Clupeidae; Homo.
REFERENCE 1 (bases 1 to 748)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strassberg, Ph.D.
Email: cga@bcr-remail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of
Orthopedics
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 499-525, >AT rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLY(A)=Yes.
FEATURES
source
Location/Qualifiers
1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EZ1-bhl-1-20-0-UI"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Left Pelvis; Vector: pTZ19-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_Ch2 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma Grade II. The library
was constructed according to Bernaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTZ19-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TGATCAGCGT.
TAG_TISSUE=grade-2-chondrosarcoma
TAG_LIB=UI-H-EZ1
TAG_SEQ=ATCTAATATG"
ORIGIN
Query Match 100.0%; Score 20; DB 13; Length 748;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTCATCTCATGTTGATG 20
|||||
DB 625 GGTCATCTCATGTTGATG 644
RESULT 29
LOCUS BG186679/c 790 bp mRNA linear EST 21-APR-2001
DEFINITION RST5654 Athersys RACE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG186679
VERSION BG186679.1 GI:13708366
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bonyfishes; Actinopterygii; Clupeiformes; Clupeidae; Homo.
REFERENCE 1 (bases 1 to 790)

AUTHORS
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Maye,R., Smith,B., Veloso,N., Klika,A., Hess,J., Cochran,K., Lo,K., Ofendbacher,J., Panzig,U. and Ducar,M.
Creation of genome-wide protein expression libraries using random activation of gene expression

JOURNAL
MEDLINE
21227151
11339013
Nat. Biotechnol. 19 (5), 440-445 (2001)

TITLE
Creation of genome-wide protein expression libraries using random activation of gene expression

COMMENT
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 353.
Location/Qualifiers

FEATURES
source
1..790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 100.0%; Score 20; DB 12; Length 790;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 GGTCATCTCATGTTGATG 20
|||||
255 GGTCATCTCATGTTGATG 236

Db
255 GGTCATCTCATGTTGATG 236

RESULT 30
AY414079/c 809 bp DNA linear GSS 17-DEC-2003
LOCUS
DEFINITION
Pan troglodytes XPA gene, VIRUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION
AY414079
VERSION
AY414079.1 GI:39770041
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
1 (bases 1 to 809)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
Science 302 (5652), 1960-1963 (2003)

JOURNAL
PUBMED
14671302

REFERENCE
AUTHORS
2 (bases 1 to 809)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment.
Location/Qualifiers

FEATURES
Location/Qualifiers

source
1..809
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

gene
1..809
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 100.0%; Score 20; DB 29; Length 809;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 GGTCATCTCATGTTGATG 20
|||||
743 GGTCATCTCATGTTGATG 724

Db
743 GGTCATCTCATGTTGATG 724

RESULT 31
AY414078/c 813 bp DNA linear GSS 17-DEC-2003
LOCUS
DEFINITION
Homo sapiens XPA gene, VIRUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION
AY414078
VERSION
AY414078.1 GI:39770040
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 813)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
Science 302 (5652), 1960-1963 (2003)

JOURNAL
PUBMED
14671302

REFERENCE
AUTHORS
2 (bases 1 to 813)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment.
Location/Qualifiers

FEATURES
source
1..813
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

gene
1..813
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 100.0%; Score 20; DB 29; Length 813;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 GGTCATCTCATGTTGATG 20
|||||
743 GGTCATCTCATGTTGATG 724

Db
743 GGTCATCTCATGTTGATG 724

RESULT 32
AL545429 873 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION
AL545429 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS01027YAA03 5-PRIME, mRNA sequence.
Location/Qualifiers

ACCESSION AL545429
 VERSION AL545429.2 GI:31267264
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 873)
 REFERENCE Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 AUTHORS Full-length cDNA libraries and normalization
 TITLE Unpublished (2001)
 JOURNAL On Feb 15, 2001 this sequence version replaced gi:12877910.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4910.r For more information about this cluster, see
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSOD1027AA020P1&cluster=4910.r. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CSOD1027AA020P1.
 Location/Qualifiers
 source
 1. 873
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1027AA03"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 100.0%; Score 20; DB 9; Length 873;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20
 ||||||||||||||||
 Db 809 GGTCCTACTCATGTTGATG 790

RESULT 33
 BM468458 1059 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6432269 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535843
 DEFINITION 5', mRNA Sequence.
 ACCESSION BM468458
 VERSION BM468458.1 GI:18517500
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1059)
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgarbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Place: LLNL12224 row: 5 column: 04

FEATURES
 source
 High quality sequence stop: 662.
 Location/Qualifiers
 1..1059
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5535843"
 /tissue_type="lelomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH_MGC_71"
 /note="Orogen: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN
 Query Match 100.0%; Score 20; DB 12; Length 1059;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20
 ||||||||||||||||
 Db 743 GGTCCTACTCATGTTGATG 724

RESULT 34
 AL573331 1114 bp mRNA linear EST 31-MAY-2003
 LOCUS AL573331 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CSOD1042YC03 3-PRIME, mRNA sequence.
 ACCESSION AL573331
 VERSION AL573331.2 GI:31294686
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1114)
 REFERENCE Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 AUTHORS Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 16, 2001 this sequence version replaced gi:12932473.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4910.r For more information about this cluster, see
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSOD1042AB02NP1&cluster=4910.r. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CSOD1042AB02NP1.
 Location/Qualifiers
 source
 1..1114
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1042YC03"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 100.0%; Score 20; DB 9; Length 1114;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20
 ||||||||||||||||

Db 559 GGTCCTACTCATGTTGATG 578

RESULT 35
AL548961/c 1129 bp mRNA linear EST 31-MAY-2003
LOCUS AL548961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1042YC03 5-PRIME, mRNA sequence.
ACCESSION AL548961
KEYWORDS AL548961.2 GI:31270783
SOURCE EST.
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Euteleostomi; Primates; Catarrhini; Homidae; Homo.
Li W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12884484.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4910.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1042AB02QPl&cluster=4910.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1042AB02QPl.

FEATURES
source
1..1129
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1042YC03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1129;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
|||||
Db 818 GGTCCTACTCATGTTGATG 799

RESULT 36
B0879392/c 1174 bp mRNA linear EST 16-AUG-2002
LOCUS B0879392 AGENCOURT 818253 lupski_dorsal_root_ganglion Homo sapiens cDNA
DEFINITION clone IMAGE:6181232 5', mRNA sequence.
ACCESSION B0879392
KEYWORDS B0879392.1 GI:22271400
SOURCE EST.
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Euteleostomi; Primates; Catarrhini; Homidae; Homo.
Li W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (1999)
NIH-MGC http://mgs.nci.nih.gov/
Mammalian Gene Collection (MGC)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLM13565 row: e column: 09
High quality sequence start: 85
High quality sequence stop: 369.
Location/Qualifiers
1..1174
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6181232"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCAAGGCGCGCG-3' and
5'-GACTAGTTCTAGATGCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 1174;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20
|||||
Db 242 GGTCCTACTCATGTTGATG 223

RESULT 37
AL531628 1201 bp mRNA linear EST 23-MAY-2003
LOCUS AL531628 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM002YF23 3-PRIME, mRNA sequence.
ACCESSION AL531628
KEYWORDS AL531628.2 GI:31069460
SOURCE EST.
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Euteleostomi; Primates; Catarrhini; Homidae; Homo.
Li W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:12795121.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4910.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM002C12NP1&cluster=4910.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DM002C12NP1.
Location/Qualifiers
1..1201

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDM0021P23"
/cisue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_1ib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pcwvSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pcwvSPORT 6
vector. Library was not normalized."

```

ORIGIN

```

Query Match          98.0%; Score 19.6; DB 9; Length 1201;
Best Local Similarity 95.0%; Pred. No. 1.8e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GGTCCATCTCATGTTGATG 20
|||||
563 GGTCCATCTCATGTTGATG 582

```

```

RESULT 38
LOCUS A1990924 640 bp mRNA linear EST 08-SEP-1999
DEFINITION ws24f07.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2498149 3'
similar to gb:Di4533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS
(HUMAN); mRNA sequence.
A1990924
A1990924
A1990924.1 GI:5837821
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 640)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/bbrp/image/image.html
Seq primer: -80UP from Gibco
High quality sequence stop: 459.
Location/Qualifiers
1..640

```

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2498149"
/cisue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_1ib="NCI CGAP GC6"
/notes="Vector: pTR73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI-CGAP-GC4 was prepared, and
as circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

```

ORIGIN

```

Query Match          95.0%; Score 19; DB 9; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2 GTCCATCTCATGTTGATG 20
|||||
605 GTCCATCTCATGTTGATG 623

```

RESULT 39

```

LOCUS BE763751 467 bp mRNA linear EST 19-SEP-2000
DEFINITION RCS-NT0053-140600-022-D04 NT0053 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE763751
VERSION BE763751.1 GI:10193675
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 467)

```

REFERENCE

```

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Negal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

```

TITLE

```

JOURNAL Laboratory of Cancer Genetics
MEDLINE Ludwig Institute for Cancer Research
PUBMED Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Contact: Simpson A.J.G.
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2-RCS-NT0053-140
600-022-D04et3=2000-06-14et4=1)
Seq primer: puc 18 forward
High quality sequence start: 3
High quality sequence stop: 464.
Location/Qualifiers
1..467

```

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1ib="NT0053"
/notes="Organ: nervous tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORSRES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

ORIGIN

```

Query Match          92.0%; Score 18.4; DB 10; Length 467;
Best Local Similarity 95.0%; Pred. No. 4.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 GGTCCATCTCATGTTGATG 20
|||||
416 GGTCCATCTCATGTTGATG 435

```

DB

RESULT 40
 B0321319/c
 LOCUS B0321319 568 bp mRNA linear EST 17-MAY-2002
 DEFINITION QV4-CT0491-080800-341-a11 CT0491 Homo sapiens cDNA, mRNA sequence.
 ACCESSION B0321319
 VERSION B0321319.1 GI:20929303
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 568)
 Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.U., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL 20202663
 MEDLINE 10737800
 PUBMED
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV4-CT0491-
 080800-341-a11&t3=2000-08-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 45.
 Location/Qualifiers
 1..568
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CT0491"
 /note="Organ: colon; Vector: puc18; Site 1: Sma1; Site 2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No.
 196,716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 13; Length 568;
 Best Local Similarity 95.0%; Pred. No. 4.7e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTGTGATG 20
 |||||
 DB 566 GGTCCATCTCGTGTGATG 547
 |||||
 RESULT 41
 B057361
 LOCUS B057361 625 bp mRNA linear EST 29-SEP-2003
 DEFINITION B057361 NIBB Mochii normalized Xenopus laevis cDNA, mRNA sequence.
 ACCESSION B057361
 VERSION B057361.1 GI:17470155
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 625)
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and
 Kohara, Y.
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 JOURNAL
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp
 The information of this clone is available through the following
 URL.
 http://xenopus.nibb.ac.jp.
 Location/Qualifiers
 1..625
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL034n06"
 /tissue_type="whole embryo"
 /dev_stage="stage 15"
 /clone_lib="NIBB Mochii normalized Xenopus neurula
 library"
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 12; Length 625;
 Best Local Similarity 95.0%; Pred. No. 4.9e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGTCCATCTCATGTGTGATG 20
 |||||
 DB 49 GGTCCATCTCATGTGTGATG 68
 |||||
 RESULT 42
 A1634902
 LOCUS A1634902 711 bp mRNA linear EST 26-APR-1999
 DEFINITION t274c02.x1 NCI CGAP Panel Homo sapiens cDNA clone IMAGE:2294306 3'
 similar to gb:U14533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS
 (HUMAN); mRNA sequence.
 ACCESSION A1634902
 VERSION A1634902.1 GI:4686232
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 711)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Life Technologies catalog #: 11548-013
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E.B. Consortium/ILMB at:
 www-bio.liml.gov/bbrp/image/image.html
 Seq primer: -40UP from Glibco
 High quality sequence stop: 409.
 Location/Qualifiers
 1..711
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2294306"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"

/clone.lib="NCI CGAP Pan1"
 /note="Organ: pancreas; Vector: PCMV-SPORE; Site: 1; Sal1;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.72 kb. Life Technologies catalog #: 11546-013"

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 711;
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 GGTCCTACTCATGTTGATG 20
 |||||
 180 GGCCCTACTCATGTTGATG 199

Db

RESULT 43
 BG187563 854 bp mRNA linear EST 21-APR-2001
 LOCUS RST6691 Atherys RAGE library Homo sapiens cDNA, mRNA sequence.
 DEFINITION BG187563
 ACCESSION BG187563.1 GI:13709378
 VERSION EST.
 KEYWORDS
 ORGANISM Homo sapiens (human)
 SOURCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,
 Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
 Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
 Maye, R., Smith, E., Veloso, N., Kliska, A., Hess, J., Cothren, K., Lo, K.,
 Offenbacher, J., Danzig, J. and Ducar, M.
 Creation of genome-wide protein expression libraries using random
 activation of gene expression

JOURNAL MEDLINE PUBMED

11329013
 Contact: Scott J. Cain
 Atherys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atherys.com

High quality sequence stop: 482.

FEATURES

source

1..854
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone.lib="Atherys RAGE library"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 854;
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 GGTCCTACTCATGTTGATG 20
 |||||
 602 GGTCCTACTCATGTTGATG 621

Db

RESULT 44
 AL571156 857 bp mRNA linear EST 31-MAY-2003
 LOCUS AL571156
 DEFINITION AL571156 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CS0D10277AA03 3-PRIME, mRNA sequence.
 AL571156
 AL571156.2 GI:1292558
 EST.

ORIGIN

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 16, 2001 this sequence version replaced gi:12928170.
 Contact: Genoscope

JOURNAL

Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4910.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D10277AA02NP1&cluster=4910.r. Contact :
 Peng Liang Email : fliang@life.com URL : Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D10277AA02NP1.

FEATURES

source

1..857
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D10277AA03"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone.lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 857;
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 GGTCCTACTCATGTTGATG 20
 |||||
 586 GGTCCTACTCATGTTGATG 607

Db

RESULT 45
 CC448351 1024 bp DNA linear GSS 22-MAY-2003
 LOCUS ZMMBBC0329P17r ZMMBBC Zea mays subsp. mays genomic clone
 DEFINITION ZMMBBC0329P17 3', genomic survey sequence.
 ACCESSION CC448351
 VERSION CC448351.1 GI:30995914
 KEYWORDS GSS.

Ze mays subsp. mays (maize)
 Zea mays subsp. mays

ORIGIN

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
 Rouzard, K., Fuks, G., Yu, Y., Wing, R. and Messing, J.
 Sequencing of the maize genome at PCR (2003b)
 Unpublished (2003)

JOURNAL

Dr. Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735

Email: bharti@wakeman.rutgers.edu
 Seq primer: SP6
 Class: BAC ends
 High quality sequence start: 66.
 Location/Qualifiers

FEATURES

source

1..1024
 /organism="Zea mays subsp. mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /sub_species="mays"
 /db_xref="taxon:4578"
 /clone="ZMMBC0329P17"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMMBRC"
 /note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN

Query Match 92.0%; Score 18.4; DB 28; Length 1024;
 Best Local Similarity 95.0%; Pred. No. 6e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 GGTCCATCTCATGTTGATG 20
 |||||
 759 GGTCCATCTCATGTTGATG 778

RESULT 46

AM646323 420 bp mRNA linear EST 26-APR-2001
 LOCUS cm63e10.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
 DEFINITION laevis cDNA clone PBX0162E10 5', mRNA sequence.
 ACCESSION AM646323
 VERSION AM646323.1 GI:7403809
 KEYWORDS EST.

ORGANISM

Xenopus laevis (African clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 420)

REFERENCE

AUTHORS

Blackshear,P.J., Lai,W.S., Thorn,J.M., Kennington,E.A., Staffa,N.G.,
 Jr., Moore,D.T., Bouffard,G.G., Beckstrom-Sternberg,S.M.,
 Touchman,J.W., Donald,M.F. and Soares,M.B.
 The NIH8S Xenopus maternal EST project: interim analysis of the
 first 13,879 ESTs from unfertilized eggs
 Gene 267 (1), 71-87 (2001)

TITLE

JOURNAL MEDLINE PUBMED

COMMENT

Contact: Perry J. Blackshear
 Office of Clinical Research and Laboratory of Signal Transduction
 National Institute of Environmental Health Sciences
 A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
 USA
 Tel: 919 541-4899
 Fax: 919 541-4571
 Email: black009@niehs.nih.gov
 Clone is available through Research Genetics, Inc., 2130 Memorial
 Parkway, Huntsville, AL 35901
 phone 800-533-4363 ext.cdna, fax 256-536-9016 att.cdna, email
 cdna@resgen.com
 DNA Sequencing and analyses performed by National Institutes of
 Health Intramural Sequencing Center (NISC).

FEATURES

source

1..420
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="PBX0162E10"

/sex="female"
 /tissue_type="unfertilized egg"
 /cell_type="unfertilized egg"
 /dev_stage="unfertilized egg"
 /lab_host="DH10B"
 /clone_lib="Blackshear/Soares normalized Xenopus egg
 library"

/note="Vector: pRTT3-Pac; Site 1: EcoRI; Site 2: NotI;
 PolyA-selected mRNA was prepared from unfertilized Xenopus
 laevis eggs. The library was constructed in the vector
 pRTT3-Pac as described in Bonaldo, M.F., Lennon, G. and
 Soares, M.B. 'Normalization and subtraction: two
 approaches to facilitate gene discovery'. Genome Research
 6:791-806, 1996. The first strand synthesis used a
 NotI-dT18 primer; double stranded cDNAs were ligated to
 EcoRI adaptors, digested with NotI, and directionally
 cloned into the NotI and EcoRI-digested pRTT3-Pac vector.
 The library contained approximately 7.2 x 10⁵
 recombinants, with average insert sizes of 1-1.5 kb."

ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 420;
 Best Local Similarity 94.7%; Pred. No. 1.2e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

2 GTCCATCTCATGTTGATG 20
 |||||
 202 GTCCATCTCATGTTGATG 220

RESULT 47

BJ083380 492 bp mRNA linear EST 29-SEP-2003
 LOCUS BJ083380 NIBB Mochii normalized Xenopus tailbud library Xenopus
 DEFINITION laevis cDNA clone XL085a05 3', mRNA sequence.
 ACCESSION BJ083380
 VERSION BJ083380.1 GI:17578422
 KEYWORDS EST.

ORGANISM

Xenopus laevis (African clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 492)

REFERENCE

AUTHORS

Kiteyama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.

Expressed genes in X. laevis embryo

Unpublished (2001)

Contact: Tadasu Shin-i

Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yatae, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tehni@genes.nig.ac.jp

The information of this clone is available through the following
 URL:
 http://xenopus.nibb.ac.jp.

FEATURES

source

1..492
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL085a05"
 /tissue_type="whole embryo"
 /dev_stage="stage 25"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud
 library"

ORIGIN

Query Match 87.0%; Score 17.4; DB 12; Length 492;
 Best Local Similarity 90.0%; Pred. No. 1.3e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGTCATCTCATGTTGATG 20
 |||||
 Db 429 GGTCATCTCATGTTGATG 448

RESULT 48
 LOCUS AL637062/c 587 bp mRNA linear EST 07-NOV-2003
 DEFINITION AL637062 XCC-neurula silurana tropicalis cDNA clone TNeu03d20 5',
 mRNA sequence.

ACCESSION AL637062
 VERSION AL637062.2 GI:38215704
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 587)
 TITLE Croling, M.D.R., Ahurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
 JOURNAL Sanger Xenopus tropicalis EST project 2001 (11_2003)
 COMMENT Unpublished (2003)
 On Nov 7, 2001 this sequence version replaced gi:16789041.
 Contact: Huckle B

FEATURES
 source
 location/Qualifiers
 1..587
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TNeu03d20"
 /dev_stage="neurula"
 /lab_host="Escherichia coli DH10B"
 /clone_lib="XCC-neurula"
 /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dt primed from 5ug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 587;
 Best Local Similarity 94.7%; Pred. No. 1.4e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GTCCATCTCATGTTGATG 20
 |||||
 Db 507 GTCCATCTCATGTTGATG 489

RESULT 49
 LOCUS CD434346 828 bp mRNA linear EST 03-JUN-2003
 DEFINITION EL01N0333A02.b Endosperm_3 Zea mays cDNA, mRNA sequence.
 ACCESSION CD434346
 VERSION CD434346.1 GI:31349989
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

REFERENCE
 AUTHORS clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 828)
 TITLE lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
 Messing, J.
 JOURNAL Sequencing of the maize endosperm ESTs
 COMMENT Unpublished (2002)
 Contact: lai, jinsheg
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T3.

FEATURES
 source
 location/Qualifiers
 1..828
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="W22"
 /db_xref="taxon:4577"
 /tissue_type="Endosperm of 7-23DAP"
 /clone_lib="Endosperm_3"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 828;
 Best Local Similarity 94.7%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GTCCATCTCATGTTGATG 20
 |||||
 Db 296 GTCCATCTCATGTTGATG 314

RESULT 50
 LOCUS B1261004 808 bp mRNA linear EST 17-JUL-2001
 DEFINITION 60297217F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:511539 5',
 mRNA sequence.

ACCESSION B1261004
 VERSION B1261004.1 GI:14819850
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 808)
 TITLE NIH-MGC http://mgc.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLM at:
 http://image.llnl.gov
 Plate: L1M611270 row: j column: 20
 High quality sequence start: 63
 High quality sequence stop: 173.
 location/Qualifiers
 1..808
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:511539"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;

Site 2: Sail; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 85.0%; Score 17; DB 12; Length 808;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GTCCATACCTCATGTTGA 18
Db 48 GTCCATACCTCATGTTGA 64

Search completed: August 17, 2004, 15:45:29
Job time : 2606 secs

THIS PAGE BLANK (USPTO)